NASA CR-
141789

(NASA-CR-141789) POST-TEST NAVIGATION DATA ANALYSIS TECHNIQUES FOR THE SHUTTLE ALT Final Report (Intermetrics, Inc.) 96 p HC $4.75 CSCL 17G

N75-23629

Unclas
G3/13 20788

INTERMETRICS

POST-TEST NAVIGATION DATA ANALYSIS TECHNIQUES FOR THE SHUTTLE ALT

IR-109-75

15 APRIL 1975

INTERMETRICS INCORPORATED • 701 CONCORD AVENUE • CAMBRIDGE, MASSACHUSETTS 02138 • (617) 661-1840

CONTENTS

CONTENTS (Continued)

1. INTRODUCTION

1.1 Objective of This Document

The Shuttle Approach and Landing Tests (ALT) will provide a large quantity of data on the shuttle navigation system capabilities and performance. It is not practical to assume that all aspects of the navigation system can be tested during the flight itself. For example, it is expected that many features of the redundancy management system will not be exercised during the ALT. The use of effective post-test processing of recorded navigation sensor data can greatly increase the information return from the ALT and from all tests that follow by providing the capability for evaluating the existing navigation system and for exercising and evaluating alternative versions of the navigation system. This can be done without altering or increasing the complexity of the actual tests. It is the purpose of this document to define cost-effective approaches to such post-test data analysis.

One of the design goals chosen for the development of the post-test analysis techniques should be to consider approaches that are useful both for the ALT and for the mission data analysis of the shuttle after it has become operational. This will greatly aid the performance evaluation and design-modification processes during the early operational phases of the shuttle program. In this manner, tools developed for the ALT data analysis will become more cost-effective by way of maximizing their usefulness in other phases of the shuttle development.

The post-test data analysis activities, as perceived in this document, are categorized as follows:

Quick-Look Processing: Quick-look processing (QLP), as the the term implies, is the preliminary, rapid processing of recorded data. This includes such activities as data listing and plotting, sensor status reporting, failure analysis, and navigation mode monitoring throughout the flight. This analysis category allows for a rapid determination of test performance as it isolates the major problem areas and other unusual events.

Post-Flight Navigation Processing: Post-flight navigation processing (PFNP) refers herein to the execution of the navigation functions through the use of flight recorded data. Post-flight versions of the operational filters, navigation functions, and redundancy management functions are used to recreate the navigation portion of the test. By altering data and/or navigation software formulations, evaluation of all aspects of the navigation software can be tested in an operational environment. In addition, alternative formulations can be tested.

Error Isolation Processing: Error isolation processing (EIP) refers to the determination of the magnitudes of those error sources that contribute to navigation error. By the isolation of the error contributors to a component level, problem areas can be identified. In addition, the determination of component errors can greatly aid in the detailed evaluation of the redundancy management formulation.

Each of these three areas of post-test data processing is described in this report. Included in the descriptions are their respective advantages, disadvantages, uses, design, operation, and data requirements.

1.2 Format of the Report

In setting forth the post-test data analysis techniques for ALT navigation data, the overall post-test processor requirements are described first. In that requirements section, operational and design requirements are discussed separately, followed by the data input requirements and the requirements on the software test.

In Section 3, the post-test data processing is described in three segments based on the natural test sequence: from quick-look analysis, to post-flight navigation processing, to error isolation processing. In each of the three segments, the level of detail includes a functional requirements description and a description of the software requirements. Section 3 concludes with a summary, emphasizing the tradeoffs that must remain open and subject to analysis until final definition has been achieved in the shuttle data processing system and the overall ALT plan.

A development plan for the implementation of the ALT post-test navigation data processing is presented in Section 4. This plan is a necessary adjunct of this report so as to show the feasibility of having the system as described herein completed in time to support the shuttle approach and landing tests. The document closes with a summary and conclusions.

2. POST-TEST NAVIGATION SYSTEM REQUIREMENTS

2.1 Operational Requirements

The ALT phase of the Orbiter development will qualify the vehicle for the last and most critical phase of its operational use, namely, approach and landing. To accomplish this test, the Orbiter will be carried aloft piggy-back on a modified Boeing 747 for one or more captive flights. Following these captive flights, the Orbiter will be carried aloft and released for a free-fall duplication of the approach and landing.

Because of the cost of each flight, in terms of support personnel and systems, the ALT will place a premium on rapid and efficient attainment of test objectives. All design features necessary should be incorporated in the post-test data processing to ensure that all possible information about the navigation system performance is extracted from each flight, and that no flight test or flight test phase needs to be duplicated because of a data processing or analysis deficiency, whether in concept or in implementation.

Because of the need to optimize carefully both the status of the Orbiter systems after each flight and the flight test plan itself, it is not anticipated that fast turnaround between individual flights will be the rule. Nevertheless, because of the need to identify immediately the sources of anomalies, discrepancies, and failures, it is anticipated that severe schedule pressures will be applied to the test data analysis operation following each flight. Following anomaly source identifications, it will often be necessary to require hardware or software modifications under conditions of limited time and approaching deadlines. Original vendors will be called in, and there will be redesign, simulation, and repeats of qualification tests. All these activities place tremendous pressures on early and efficient problem diagnosis.

In the light of these and other considerations, the next subsection discusses the design requirements for the post-ALT data analysis.

2.2 Design Requirements

2.2.1 General

The navigation data analysis system must have a clear, simple method for handling all parameters that might vary during the entire ALT phase. This means that measurement lists, gains, biases, units, must all be alterable under proper project control. Alteration must include timely specification in the flight test plan, loading, verification, and documentation. The software system must be able to run in a pre-flight checkout mode with

INTERMETRICS INCORPORATED • 701 CONCORD AVENUE • CAMBRIDGE, MASSACHUSETTS 02138 • (617) 661-1840

sample or previous test data so as to provide an unambiguous "ready" status in support of the flight.

2.2.2 Front End

The front end of the system should be a central file maintenance module. As input, this activity should be able to accept tapes or other machine-readable forms from as many different data sources as possible, with the main design emphasis being placed upon telemetry and onboard recordings from the Orbiter data processing system (DPS). Additionally, the maintenance module must be able to accommodate tapes containing downlinked data from the telemetry station, if any; tapes from the acceptance checkout equipment (ACE) system in the hangar; and tapes from the ground-based tracking organization(s).

The primary design aim of the central file maintenance module is to place data of various formats into a data base that would be accessible by numerous analysis processes. There is a strong design requirement to develop those analysis processes into a common higher-order language, as a matter of cost efficiency in program development.

The maintenance module should be self-reporting and self-documenting in its performance. It should output a post-run analysis that will help with data control and analysis in the succeeding phases. Specifically, the post-run output should describe the input tape and format, the output tapes and format, the anomalies, gaps, etc., discovered, and the clear failures discovered on a first-run basis.

2.2.3 Blanking and Blending

Another front end-function involves the ability to prepare pseudo-flight tapes for the purpose of simulating the presence or absence of selected data categories. This capability would enable study of alternative test flight scenarios when a major sensor item has failed, is restored, or has been modified. The inputs to the blanking and blending module would contain the preprocessed flight tapes from the maintenance module; simulated time histories of sensor data; and miscellaneous instructions. The output would be a new, pseudo-flight tape containing modified time histories. This new output tape can be written:

a. in the same standard format as the input tape if substitutions only are involved;

b. in longer standard format if supplemental time histories are added; and

c. in a shorter standard format if there is blanking of certain parameters.

As in the case of the maintenance module, the blanking and blending module should contain a self-documenting listing of the input or source tapes, and the modifications to be accomplished.

2.2.4 Specific Processor Design Requirements

The post processors must be able to handle data over the entire operational range of the shuttle ALT. The processors must also be simple to use and modify and must have sufficient commonality with the operational software (variable names, algorithm formulation, etc.) to allow easy modification of the operational software based on post-test results. The following text presents specific constraints that must be placed on the post-test software.

a. The data analysis processors must be able to handle data obtained from the following ALT mission characteristics:

(1) Maximum shuttle speed - 1,600 fps

(2) Maximum shuttle altitude - 70,000 ft

(3) Maximum shuttle load factor - 2 g

These constraints have been defined to assume adequate post-test processor algorithm formulation and precision.

b. The post-test processors must be capable of simultaneously utilizing the data from the following sensors:

(1) 3 Inertial Measurement Units

(2) 3 TACAN receivers

(3) 3 MSBLS receivers

(4) 3 Air Data Units

(5) 2 Radar Altimeters

(6) Theodolite Ground Tracking Data

(7) C-Band Radar Ground Tracking Data

The requirement for simultaneous use of this data in the post-test processors is based on the desire to use the post-flight processor (see Section 3.2) as a method of evaluating redundancy management algorithms.

c. The post-test processors must be capable of processing multiple input tapes. They must also be capable of accepting flight-recorded or downlink-recorded shuttle data. This would permit maximum utilization of available data.

d. The post-test processors should be constructed in a structured top-down manner. This approach provides increased program readability and maintainability. A major goal of this approach is to minimize the time required to change portions of the post-test processors to reflect changes in output desired, candidate navigation or redundancy management algorithms to be tested, etc.

e. The post-test processors should also be constructed using the HAL language. This requirement is most critical for the post-flight processor. As envisioned (see Section 3.2), the post-flight processor contains a post-flight version of the operational navigation and redundancy management algorithms. The use of the operational computer language, HAL, would facilitate verification of the similarity between operational and post-flight versions. In addition, any algorithm changes that are made and tested post-flight and are to be incorporated into the operational version can be made most accurately by using the same computer language for the operational and post-flight computer programs.

2.3 Data Collection Requirements

Three major types of data, and a number of minor types, must be collected for the design as presently envisioned. The major types are the preset navigation parameters and constants, the shuttle navigation sensor data recorded throughout the flight, and the ground tracking trajectory data. This section presents a preliminary list of variables for each of the major data types. The few minor types are introduced within the functional descriptions found in Section 3.

a. The preset navigation parameters and constants are time invariant variables such as earth rate, conversion factors, pad loaded constants, etc. Many of these parameters are known apriori, but others must be recorded at the time of the pad load to ensure an exact duplication of the on-board system. A list of these parameters (which will be used to generate a parameter file tape; see Section 3.2) is presented in Table 2-1*.

*The data lists presented in Tables 2-1 and 2-2 are based on variables presented in the document "Approach and Landing Test Level C Requirements for Navigation", Sept. 3, 1974.

Table 2-1

Present Navigation Parameters and Constants

Variable Name	Data Type	Variable Description
ALT_MLSAZ ALT_MLSR ALT_MLSEL ALT_TACAN	Scalar Scalar Scalar Scalar	Altitude above ref. ellipsoid of MLS azimuth, elevation, and range radars and the TACAN station, respectively
AZ_RADAR_BEARING	Scalar	Bearing from true north of MLS azimuth radar boresight axis
BARO_DATA_GOOD	Bit	Bit set at start of measurement processing sequence to indicate presence of good baro altimeter data
BIASACC1(2,3)	Vec(3)	Acceleration bias vectors for the 3 states
BIAS_ALIGN1(2,3)	Vec(3)	IMU misalignment biases for the 3 states
BIAS_AZMLS BIAS_BARO BIAS_DME BIAS_ELMLS BIAS_MLSRANGE BIAS_VOR	Scalar Scalar Scalar Scalar Scalar Scalar	Initial values of the sensor biases for MLS measurements and TACAN measurements and BARO measurement
DELR_RAMP	Scalar	Distance over which velocity weighting of radar altimeter is ramped from 0 to its maximum value
e	Scalar	Constant natural log

Table 2-1 (Continued)

Variable	Data Type	Variable Description
ELLIPT	Scalar	Earth Ellipticity constant
EL_RADAR_BEARING	Scalar	Bearing from true north of MLS elevation radar boresight axis
H_FBAR	Scalar	Altitude below which baro is inhibited. Automatic mode
H_INBAR	Scalar	Altitude at which baro measurements are initiated. Automatic mode
H_RADALT	Scalar	Altitude below which radar altimeter is initiated
J2	Scalar	Value of J2 constant in earth's gravity model
K_RES_EDIT	Scalar	Scale factor on filter mean square residual used in filter residual edit test
K_UND_WGT	Scalar	Measurement underweighting factor. Multiplies BT_E_B_ which is then added to VAR.
K_VAR_AZMLS K_VAR_ELMLS K_VAR_RMLS K_VAR_RADALT	Scalar Scalar Scalar Scalar	Scale factors on respective sensor variances used in LAND_FILTER residual edit test.

Table 2-1 (Continued)

Variable Name	Data Type	Variable Description
LAT_MLSAZ LAT_MLSEL LAT_MLSR	Scalar Scalar Scalar	Geodetic latitude of appropriate MLS sensor radar location
LAT_TACAN	Scalar	Geodetic latitude of TACAN radar location
LONG_MLSAZ LONG_MLSEL LONG_MLSR	Scalar Scalar Scalar	Longitude of appropriate MLS radar location
LONG_TACAN	Scalar	Longitude of TACAN radar location
R_GO_RAMP	Scalar	Value of slant range between shuttle and landing site below which radar altimeter data is processed in the filter
R_NAVBASE_BODYFR	Vec(3)	Navigation base location in body coordinate system
R_RADALT_BODYFR	Vec(3)	Radar altimeter location in body coordinate system
R_TACAN_EF	Vec(3)	TACAN position vector
R_TACANT_BODYF	Vec(3)	Shuttle TACAN receiver location in body coordinate system
R_MLS_ANT_BODYFR	Vec(3)	Shuttle MLS receiver location in body coordinate system
R_RMLS_EF	Vec(3)	MLS range radar position vector

Table 2-1 (Continued)

Variable Name	Data Type	Variable Description
RADALT_BIAS	Scalar	Radar altimeter bias
RGO_AZMLS_ZERO RGO_ELMLS_ZERO	Scalar Scalar	Fixed weighting vector range scale factors for MLS azimuth and elevation measurements respectively
RE	Scalar	Earth equatorial radius
RHO_SEA_LEVEL	Scalar	Sea level atmospheric density
R_RDAR_BEARING	Scalar	Bearing from ture north of MLS range radar boresight axis
r_GRAV	Scalar	Earth radius used for gravity oblateness term
SCALE_HGT	Scalar	Alatitude scale height for exponential model of atmospheric density
VAR_UNMOD_ACC_ΔT	Scalar	Variance of unmodeled acceleration times time step. Used in calculation of process noise matrix for velocity error due to platform misalignment
VAR_ACC_QUANT	Scalar	Variance of accelerometer quantization error. Used in calculation of process noise matrix
VAR_DRIFT_DT	Scalar	Variance of error in gyro drift compensation. Used in calculating process noise matrix for platform misalignment states.

Table 2-1 (Continued)

Variable	Data Type	Variable Description
VAR_AZMLS VAR_DME VAR_ELMLS VAR_H1 VAR_H2 VAR_RADALT VAR_RMLS VAR_VOR	Scalar Scalar Scalar Scalar Scalar Scalar Scalar Scalar	A-priori sensor measurement error variances (VAR_H1 and VAR_H2 used for BARO measurements).
V_RHI V_RLO	Scalar Scalar	Upper and lower limits respectively of relating wind velocity between which BARO updating is prohibited.
ε_{AZ} ε_{EL} ε_{R}	Scalar Scalar Scalar	Difference criteria in MLS range, azimuth and elevation measurement change check, respectively
ε_{TAG}	Scalar	Difference criteria used to determine if state vectors should be extrapolated (from T_FILT to T_SENSORS) to compute measurement residuals
ε_{T_MLS}	Scalar	Difference criteria to force MLS azimuth or elevation measurement incorporation even though the measurement is within ε_{AZ} or ε_{EL} of the previous measurement
μ	Scalar	Universal gravitational constant for the earth

Table 2-1 (Continued)

Variable Name	Date Type	Variable Description
π	Scalar	PI
σ_ASCENT	Vec(12)	Represents the one-sigma values of the initial filter covariance matrix diagonal elements. In local vertical coordinate frame (u,v,w); used in ASNT_INITIALIZE
σ_ACC	Scalar	RMS acceleration bias for each of the three filter states. σ_ACC same for each of the three acceleration biases of each state
σ_AZMLS	Scalar	RMS values of particular sensor measurement bias
σ_DME	Scalar	
σ_ELMLS	Scalar	
σ_RMLS	Scalar	
σ_VOR	Scalar	
σ_BARO	Scalar	Altitude dependent RMS value for BARO measurement
σ_BARO_ZERO	Scalar	RMS BARO altimeter error at zero altitude
σ_UND_WGT	Scalar	Measurement underweighting criterion

Table 2-1 (Continued)

Variable Name	Data Type	Variable Description
τ_ACC	Scalar	Acceleration measurement
τ_AZMLS	Scalar	MLS azimuth measurement
τ_BARO	Scalar	Baro measurement
τ_DME	Scalar	TACAN DME measurement
τ_ELMLS	Scalar	MLS elevation measurement
τ_RMLS	Scalar	MLS range measurement
τ_VOR	Scalar	TACAN VOR measurement
ω_ALT_RWTD ω_ALTDOT	Scalar	Prestored gain for radar altimeter measurements in the altitude and altitude rate channel (max value of altitude rate gain). Landing phase
ω_ALTDOT_RWTD	Scalar	Value of altitude rate filter gain for radar altimeter measurements in landing phase
ωE	Scalar	Earth rotation rate
ω_LAND_AZMLS ω_LAND_ELMLS ω_LAND_RMLS	Vec(6) Vec(6) Vec(6)	Prestored gain vectors for MLS azimuth, elevation, and range measurements respectively in landing phase filter. (Position and velocity updating) in the respective radar coordinate systems

b. The shuttle navigation sensor data includes the navigation parameters that must be recorded throughout the flight. This includes sensor outputs with time tags, mode status switches, data good flags, and selected computed navigation parameters like the shuttle position vector from the navigation state selection filter. A preliminary list of those parameters required for post-flight processing is presented in Table 2-2. The assumption is made that these parameters are recorded at a rate of ΔT_AVG_G_ENTRY.

c. The ground tracking trajectory data is an independent trajectory as determined from C-band and/or cinetheodolite data with associated time reference. It is assumed that that data will be presented in a processed form.

The lists of data collection requirements presented in this section define a baseline for the presently-envisioned post-test processors. Variations in this list will occur as the processor design develops further. In the meantime, this list will be useful as a preliminary definition of the data recording requirements.

Table 2-2

Shuttle Navigation Sensor Data

Variable Name	Data Type	Variable Description
AZMLS_EDIT	Bit	On-indicates MLS azimuth measurement has failed filter residual edit test
AZMLS_EDIT_OVERRIDE	Bit	Allows manual override to force MLS azimuth measurement incorporation even though filter residual edit test was not passed
AZMLS_READ	Scalar	MLS azimuth sensor reading
BARO_MODEL_OF_THE_MONTH	Bit	If on, current nav panel BARO model is used in the filter
		Data parity bit
		INS attitude from INS
		Airdata pressure
		Temp measurement
		Clock time for sync with ground track
BARO_EDIT	Bit	On-indicates BARO measurement has failed filter residual edit test
BARO_EDIT_OVERRIDE	Bit	Allows manual override of filter residual edit to force incorporation of BARO data
BARO_READ	Scalar	BARO sensor reading

Table 2-2 (Continued)

Variable Name	Data Type	Variable Description
DG_AZMLS DG_BARO DG_DME DG_ELMLS DG_RADALT DG_RMLS DG_VOR	Bit Bit Bit Bit Bit Bit Bit	Data good discretes from sensor reads
DME_EDIT	Bit	On-indicates DME measurement failed filter residual edit test
DME_EDIT_OVERRIDE	Bit	Switch to force TACAN range measurement incorporation independent of filter residual edit test
DME_READ	Scalar	TACAN DME sensor reading
DO_BARO_NAV DO_MLS_NAV DO_TACAN_NAV DO_RADALT_NAV	Bit Bit Bit	Switch which indicates which sensor measurement is to be processed
EDIT_FLAG	Bit	Local variable. On-indicates that the sensor measurement being processed has failed filter residual edit test
ELMLS_EDIT	Bit	On-indicates MLS elevation measurement has failed filter residual edit
ELMLS_EDIT_OVERRIDE	Bit	Allows manual override of filter residual edit test to force incorporation of MLS elevation measurement

Table 2-2 (Continued)

Variable Name	Data Type	Variable Description
ELMLS_READ	Scalar	MLS elevation sensor reading
IMU_FAIL	Boolian	From IMU RM module. Indicates which IMU or combination of IMU's has failed
MANUAL_EDIT_ OVERRIDE	Bit	Local variable in ENTRY_ FILTER allows manual override of filter residual edit test allowing measurement incorporation
MANUAL_BARO_ PRESENT_MODE MANUAL_TAC_MLS_ PRESENT_MODE MANUAL_RADALT_ PRESENT_MODE	Bit Bit Bit Bit	Allows manual presentation of navigation sensor data for measurement processing
MANUAL_BARO MANUAL_TAC_MLS MANUAL_RADALT	Bit Boolian Bit	Switch settings on to select BARO, TACAN, or MLS, or RADAR altimeter respectively
MLS_TIME_READ	Scalar	Time associated with MLS sensor reads
MLS_ID	Scalar	Used in identifying MLS stations so that pertinent information (locations of range, az, el, radars, etc.) can be obtained from stored tables
RADALT_EDIT	Bit	On-indicates radar altimeter measurement has failed filter residual edit

Table 2-2 (Continued)

Variable Name	Data Type	Variable Description
RADALT_EDIT_OVERRIDE	Bit	Allows manual override of filter residual edit for radar altimeter measurements
RADALT_READ	Scalar	Radar altimeter sensor read
RMLS_EDIT	Bit	On-indicates MLS range measurement has failed residual edit
RMLS_EDIT_OVERRIDE	Bit	Allows manual override of filter residual edit test for MLS range measurements
RMLS_READ	Scalar	MLS range sensor read
r_MID	Vec(3)	Shuttle position vector from navigation state selection filter
r_ONE(TWO, THREE)	Vec(3)	Current shuttle position vectors (3 states)
SENS_TIME_READ	Scalar	MTU sensor read time
TACAN_ID	Scalar	Used to identify TACAN stations whose measurements are to be processed
T_BARO T_TACAN T_RADALT	Scalar Scalar Scalar	Time of BARO, TACAN (VOR, DME) and radar altimeter sensor read respectively
T_COV	Scalar	Time tag of filter covariance (also used as dummy local variable)

Table 2-2 (Continued)

Variable Name	Data Type	Variable Description
T_EF_AZMLS T_EF_ELMLS T_EF_RMLS	Scalar Scalar Scalar	Set at start of measurement processing sequence so that a consistent set of data is used in the measurement processing cycle
T_FILT	Scalar	Time associated with r_FILT1, v_FILT1, etc.
T_IMU	Scalar	Time of IMU read
T_SENSORS	Scalar	Set at start of each measurement processing sequence; sensor reads time
T_STATE	Scalar	Time associated with current shuttle state vectors, r_ONE(TWO,THREE) and v_ONE(TWO,THREE)
VOR_EDIT	Bit	ON-indicated that VOR measurement has failed filter residual edit
VOR_EDIT_OVERRIDE	Bit	Allows manual override of filter residual edit to force incorporation of VOR measurement data
v_IMU1_CURRENT (IMU2,IMU3)	Vec(3)	Currently read (T_IMU) accumulated velocity counts from each IMU
v_MID	Vec(3)	Shuttle velocity vector from navigation state selection filter. Used for guidance calculation

Table 2-2 (Continued)

Variable Name	Data Type	Variable Description
v_ONE(TWO, THREE)	Vec(3)	Current shuttle velocity vectors (3 states)
ΔT_AZMLS ΔT_ELMLS ΔT_RMLS ΔT_RAD	Scalar Scalar Scalar Scalar	Time intervals between measurements used in time scaling the weighting vectors for the appropriate sensors in the landing phase filter
ΔV_IMU1 (IMU2, IMU3)	Vec(3)	Sensed Δv from each IMU over ΔT_IMU interval
ΔT_IMU		

2.4 Software Checkout

Development and verification of post-flight processing activity requires the use of proper management organization, coordination and control techniques. Quality control measures should be enforced throughout the development effort. These controls deal specifically with test level control, data base control, test case generation, design change control, and software/hardware interface control.

The verification of computer programs is greatly enhanced by the readability characteristics of the HAL programming language and by the use of top-down programming techniques. This allows program testing to be conducted in a purposeful, methodical way, utilizing a three-level test plan:

a. Unit Design Verification (Level 0) - The objective of a Level 0 verification is to validate the lowest-level functional units. Each unit is required to compile successfully and is further validated by a desk-check of the coding. Input/output tests should be utilized, as appropriate, to verify proper control flow, numerical results, stability, convergence, scaling, and range, etc.

b. Module Design Verification (Level 1) - Level 1 tests are designed and executed as appropriate. The objective of Level 1 testing is to test related groups or strings of software units that represent a higher level of complexity. Applications functions

INTERMETRICS INCORPORATED • 701 CONCORD AVENUE • CAMBRIDGE, MASSACHUSETTS 02138 • (617) 661-1840

are integrated. These tests verify inter-unit communications and control, name scope, and parameter passage. Preplanned test cases would be used to provide a "calibration standard" aginast which results can be evaluated.

c. System Design Verification (Level 2) - Testing at Level 2 involves the processing of simulated or real flight data by the complete system. This level requires an almost totally-integrated system, interfaced with its environment. Such testing reveals errors due to conflicts introduced by data-sharing convention violations, range of input values, sequencing requirements and interphase communications and control. Generated data should be evaluated and verified through use of post-processing editing software or by visual inspection of the results.

For the error isolation processor (see Section 3.3), it is required that the simulated nav sensor error sources be contained within the estimated $\pm$ 1σ band about the values of those error sources estimated by the EIP. Depending on the trajectory, however, some of the error sources will not be observable. This will be reflected in large values for the estimated variances of those states, thus negating the value of the above suggested test criterion. It is therefore important that one sufficient set of simulated trajectories, which would make all error sources observable at one time or another, be processed in verifying the EIP.

INTERMETRICS INCORPORATED • 701 CONCORD AVENUE • CAMBRIDGE, MASSACHUSETTS 02138 • (617) 661-1840

This three-level test sequence should be followed by acceptance tests conducted at the intended residence program facility. These tests would validate the interfacing of the system with its envisioned environment, allowing the user to confirm compliance of the delivered program with its specifications.

INTERMETRICS INCORPORATED • 701 CONCORD AVENUE • CAMBRIDGE, MASSACHUSETTS 02138 • (617) 661-1840

3. POST-TEST NAVIGATION PROCESS REQUIREMENTS

The post processing of recorded navigation data can greatly improve the efficiency and information return from the shuttle approach and landing tests (ALT). To maximize this return, early consideration must be given to the types of processing desired so that the processing concepts might have an appropriate impact on the data collection processes for the tests. Toward that end, this chapter describes post-test processing to a level of detail sufficient to outline its impact on the overall shuttle program.

Three categories of post-test data processing are presented in this chapter: quick-look processing (QLP), post-flight navigation processing (PFNP), and error isolation processing (EIP). For each category, the functional requirements and the software requirements are provided.

Figure 3-1 shows an overview of the interrelationship among the three categories of test described in this document.

3.1 Quick-Look Processor (QLP)

The quick-look processor (QLP) serves two broad post-flight navigation data processing functions, each to be carried out immediately following the flight test: 1) the QLP serves as the central data distributor and in a maintenance function; and 2) it

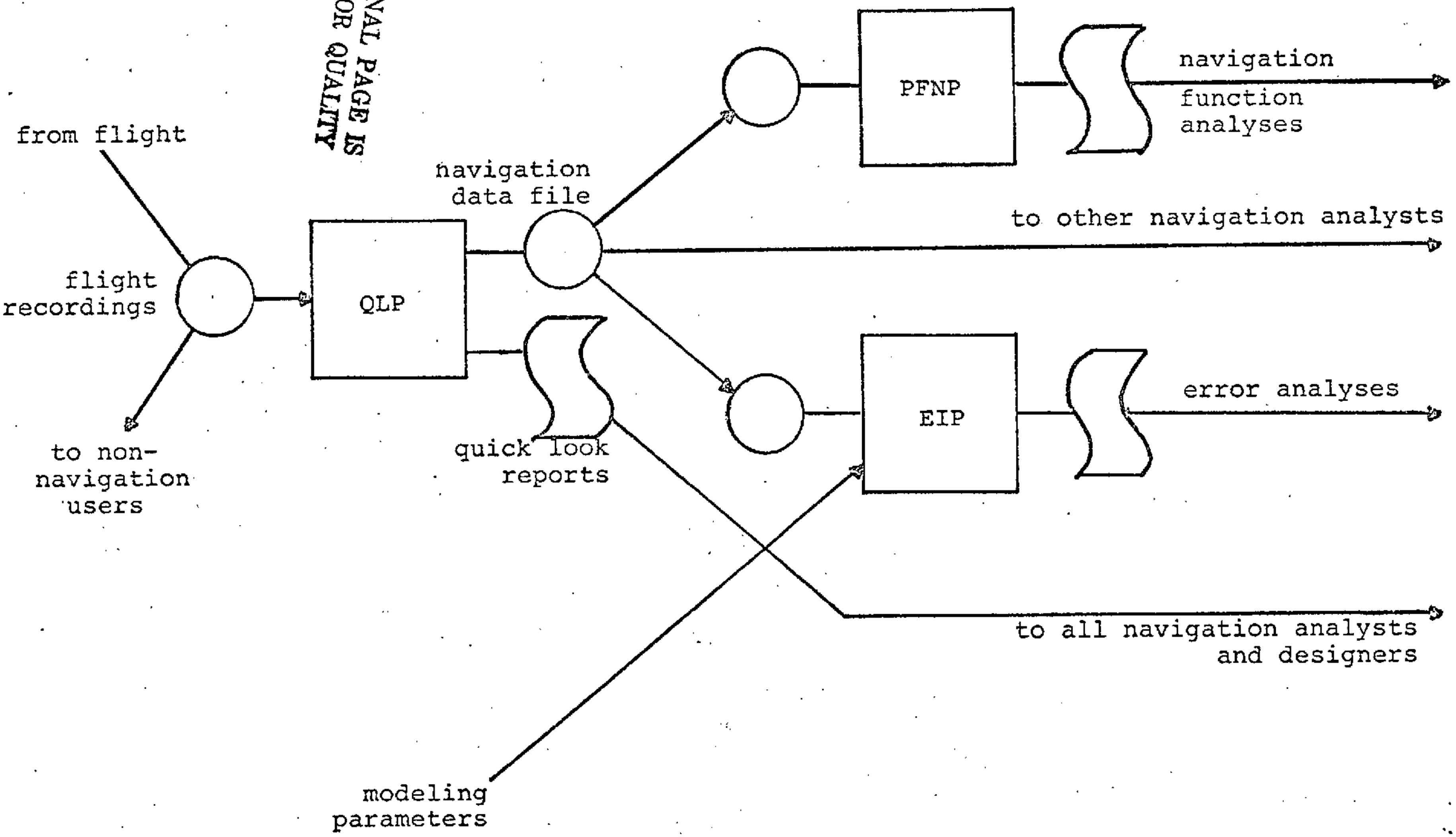


Figure 3-1 Post-Test Navigation Data Analysis Process

will create and distribute all first-cut navigation evaluation reports (i.e., the so-called 24-hour reports). A description of the activities within these two functions is presented below.

3.1.1 Functional Requirements

The inputs to the QLP are all flight navigation recordings, including: in-flight navigation data recordings, both on-board and telemetry; preset navigation parameters; and ground based tracking data. (See Section 2.3 for the detailed listing of those inputs.) The outputs from the QLP are: 1) properly formatted versions of the flight data, which will be distributed to navigation analysis and design activities, including the two described in Sections 3.2 and 3.3 herein; and 2) a series of flight navigation summary reports, which will be distributed to any organization having technical responsibility for the performance of flight navigation functions. These two output requirements place four functional requirements on the QLP, described in succeeding paragraphs.

3.1.1.1 <u>Creation and Maintenance of a Central Navigation Data File</u> - A central data maintenance function is essential to the orderly distribution of complete and readable flight data. Such a function alleviates the following potential time-consuming and costly problems for designers and analysts:

a. inability to locate data,

b. no time reference,

c. inability to synchronize two data streams,

d. improper scaling,

e. noisy data,

f. missing data points,

and a host of other difficulties. In addition, the central processor controls the amount of post-flight data processing to include only that which is necessary. It eliminates any requirement for data filing, or any other controlled handling, on the part of analysts. (This can be accomplished through the maintenance of a master copy of all flight data at the central facility.) It assures a common time frame for all data streams. It standardizes all formats. It helps to assure no duplication of effort. It avoids misplacing of data. Also, it minimizes the number of constraints on the flight data-recording equipment specifications, in that the burden of a common format is on the central facility, not on the equipment designers.

3.1.1.2 Creation and Distribution of Flight Navigation Summary Reports - The primary post-flight operational requirement is to enable the many design and analytic talents to work on necessary modifications to the navigation system as soon after the flight test as possible. This requires the immediate creation of summary

INTERMETRICS INCORPORATED • 701 CONCORD AVENUE • CAMBRIDGE, MASSACHUSETTS 02138 • (617) 661-1840

navigation reports to tag anomalies and summarize the performance of all navigation functions. And, because of the various and different perspectives of the groups concerned with navigation, there is a requirement for a number of different summary views of the flight (e.g., sensor people will be concerned with readout performance, computer people with mode switching, and so forth).

The second QLP requirements is, therefore, a set of summary reports, each with a different perspective of the flight. That set of reports represents an organized statement of the flight navigation performance.

These reports should be created within 24 hours of the flight, or at least within 24 hours of the availability of recordings, in order to expedite analysis of the flight.

3.1.1.3 Preliminary Analysis of Unusual Events - The above-mentioned summary reports will, among other things, flag unusual (unexpected or undesired) events like failures. As a third requirement on the QLP, the post-flight data processing specification should include a list of supplementary processing activities, each activity within the list dictated by the nature of a particular unusual event. For example, a particular sensor failure might dictate a processing requirement on that sensor's redundant equivalents, or on the ground tracking data, or on its calibration data, and so forth. As distinguished from the fixed summary reporting above, this QLP function exercises a set of causal requirements (that is, given an unusual event A, accomplish activity B). For this part of the QLP, only those unusual events that require immediate action should be included.

3.1.1.4 Creation of Modified Quick Looks - The outputs from the above three QLP activities will be, for every flight, to a variable extent, less than complete. There will always be unforeseen events which will not be covered by the post-flight processing specification. Nevertheless, the post-flight data processing specification must direct at least the mechanics for expediting action on such events. (The specification cannot, of course, dictate the particulars of action against unforeseen events.) For the purposes of this document, we assume the existence of such a module as the fourth requirement on the QLP. The software requirements section that follows contains little in the way of particulars, which are left as an exercise for the specification.

Note that there will be a learning curve applied to the QLP: action taken in the modified quick-look area on any flight should be covered as a specification revision for unusual event analyses on the next flight.

3.1.2 Software Requirements

The QLP routines required to accomplish the functional requirements presented in Section 3.1.1 are introduced in this subsection.

In Figure 3-2, an overview of the QLP shows it to be composed of six routines. For our purposes, herein, "routine" is not restricted to mean one distinct physical program. In fact, in most instances it is more than one. Although the function of any routine might be fixed, there is often a requirement to develop a number of programs to accomplish that function, either because the software maintenance requires a modular construction (because of ever-changing requirements), or because the physical characteristics of the various inputs (e.g., tape, disc, blocking, format) dictate several mechanical versions.

Central to the QLP is the navigation data master file (see Figure 3-2). To create that file, two routines are required: the Duplicate Flight Recordings (DFR), and the Purge, Format, and Catalog Routine (PFCR). The DFR simply creates a version of the flight test data that can be dedicated to navigation analysis alone. This assumed that navigation and non-navigation data will be found on the same flight recordings. It also assumes that the tape duplications are not accomplished by other agencies. The second routine, the PFCR, is that multi-purpose activity (probably consisting of many programs) that standardizes, formats, catalogs, sorts, and tabulates the flight data.

From the master file, four routines are required to create QLP outputs. The Duplicate Navigation Data (DND) activity acts as the central data distributor. Analysts of flight data will be

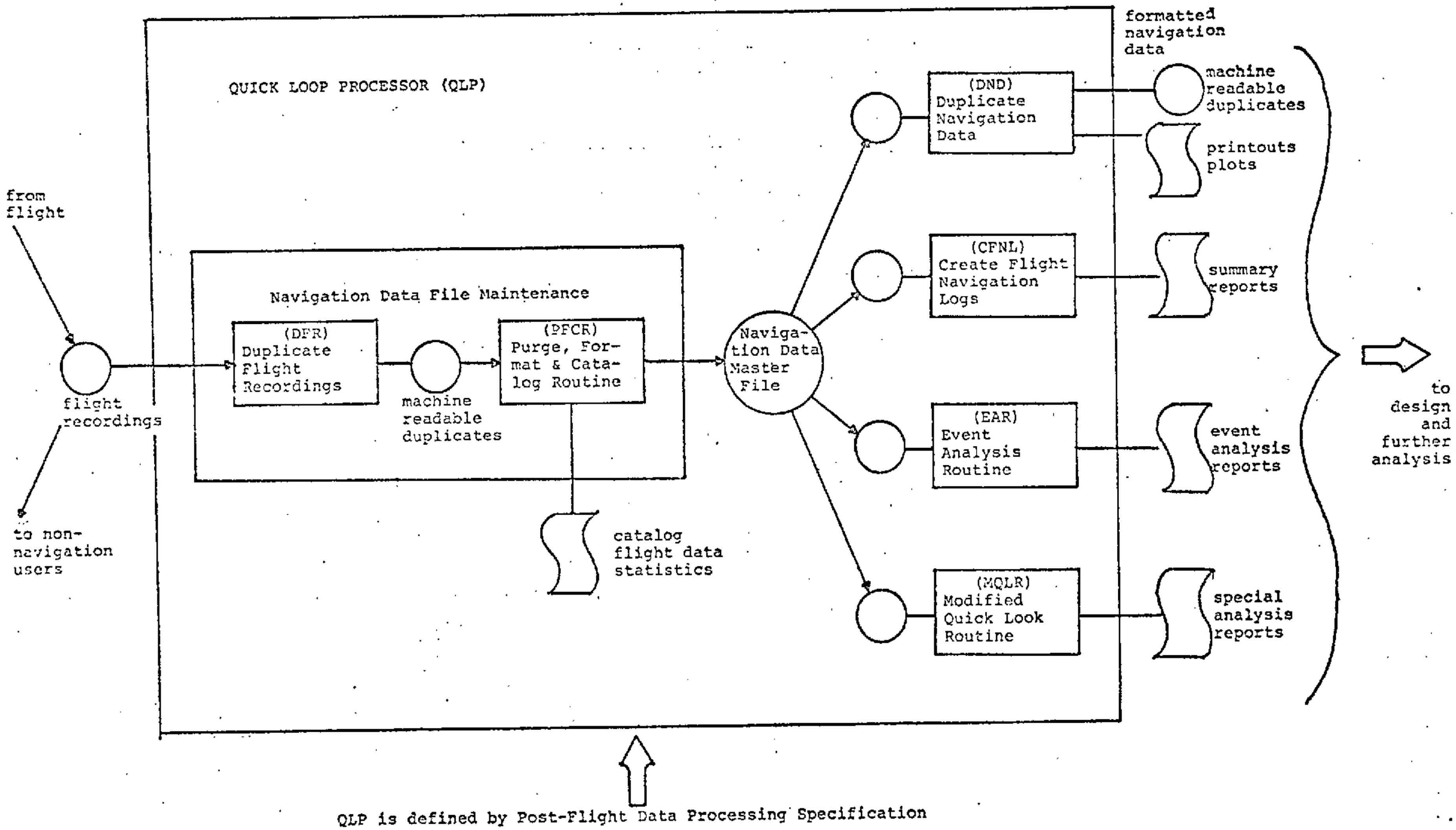


Figure 3-2 Routines Comprising the Quick-Look Processor

given data dedicated to their use alone. Therefore, they should be presented duplicates of the master file copy. The Create Flight Navigation Logs activity creates a set of summary reports, in detail sufficient for first-look analysis. The Event Analysis Routine (EAR) and Modified Quick-Look Routine (MQLR) create special analyses and reports for unusual events; the EAR for events covered by the specification, and the MQLR for those not covered by the specification.

In the following paragraphs, the software requirements for each of these six routines are covered.

3.1.2.1 Duplicate Flight Recordings (DFR) - This activity may or may not exist within the QLP, depending upon the interrelationship between the navigation post processing and non-navigation post processing activities. Assuming that navigation and non-navigation data will exist on one or more of the flight recordings, there will be a requirement for duplicating those recordings (or partitioning them, which would probably be a solution technically less practical than duplication) so as to have a dedicated version of flight data for navigation post-processing needs. A high level of planning might judiciously place that duplication function elsewhere. Until that is specified, we herein assume that the function belongs to the QLP. The software requirement is to have the capability to create "exact" duplicates of the flight recordings. For some, there may be a temptation to accomplish additional processing during this duplication process. This should not be done. Experience in most data processing activities shows that an original version of data should always be retained, because after data is processed, it is difficult to retain a picture of which portion of the processed output is raw and which has been altered. With an original on file, one can always recover that distinction when necessary. Nevertheless, any duplication process is difficult to accomplish exactly; all duplicating processes have their own noise. In any event, care should be taken to accomplish the duplication as nearly as possible.

3.1.2.2 Purge, Format, and Catalog Routine (PFCR) - This routine constitutes the most involved set of programs within the QLP, due to the large number of functions to be accomplished and to the lack of structural relationships among the functions. The varied characteristics of the inputs necessitate a "shopping list" of standardizing and cataloging processes to enable the interchangeability and cross-referencing of data necessary for the reporting processes. The list of software processing requirements is as follows:

Format Standardization: The input data is of various mechanical and logical input formats. The need to cross-reference data from the various sources dictates a standardization of format for the master file. This creates a range of process requirements including:

- Data Storage Medium Standardization - The input data probably will be physically contained in several different machine-readable forms. The PFCR should translate the data to a standard form, probably tape or disc, as dictated by the available processing equipment.
- Data Packaging - The master file is made up of n packages of data, with the word "package" used here to mean a large collection of data of a particular type (e.g., a stream of accelerometer outputs). The packaging can be accomplished logically, or physically; the choice would be dictated by tradeoffs directed toward efficient retrieval practices.
- Data Blocking - It is assumed that the processing equipment and practices dictate a standard quantum (i.e., "block") of data. A package would be partitioned into n such blocks within the PFCR.
- Data Logical Record Layout - All data should be transformed to a common "record" layout, such that a block of data equals n records. The specifics of record layout would be dictated by efficient retrieval.

- <u>Data Addressing</u> - For a tape-oriented system, data addressing is a trivial problem. For a disc-oriented system, addressing can be used for efficient retrieval. However, addressing logic development can be carried to extremes. The amount of versatility required for accessing through sophisticated addressing techniques must be dictated by the output requirements. Care should be taken not to exceed those requirements.
- <u>Data Labeling</u> - Data groups should be properly labeled; that is, all packages, all blocks, and in selected cases, some records.

<u>Data Synchronization and Calibration</u>: It is assumed that the input data will be of various units and will be referenced to several different clocks. A need for conversion to common frames of reference creates the following software requirements:

- <u>Time Conversions to a Master</u> - The different time references for each recording should be brought to a common reference if precision requirements so dictate. This can be accomplished by event commonalities between recordings, like mode switching, failures, etc., or by pre-flight clock calibrations.

INTERMETRICS INCORPORATED • 701 CONCORD AVENUE • CAMBRIDGE, MASSACHUSETTS 02138 • (617) 661-1840

- Calibration to Similar Units - The instrument data strings will have different scaling and biases (and perhaps other sensitivity variances, such as temperature). The PFCR should convert instrumentation data from various sources to the same units, to that extent dictated by the analysis requirements.
- Calibration to Similar Dimensions - Data from the different sources will be attempting identical state measurements (e.g., velocity, position) but will be of different dimensions, differing only by time derivatives or integrals. The PFCR may be the practical place to accomplish the integration or differentiation to achieve a common dimension. The alternative process location would be the analysis routines that require a common dimension. A study of the analysis specifications will dictate the choice.
- Line-up of Data Streams - It may be practical to line up data streams that fall into a particular class, e.g., all gyros, all velocity measurers, etc. The PFCR would be the practical place to accomplish this.
- Bit Stream Conversion - When data is in the form of bit streams, it may be advantageous to convert those bit streams to strings of digital bit counts.

Data Filtering: Data streams will often contain undesirable or unnecessary components. The PFCR should be assigned the task of removing those anomalies. The following requirements are placed on the software:

- Data Roundoff - Where large packages of data contain either unnecessary or meaningless digits, there will be an incentive to round-off to the most significant data.
- Data Sample - When it is reasonably certain that the data is of much finer grain than necessary, cost savings can be achieved in analysis by sampling at less frequent intervals, or by summarizing data within intervals.
- Noise Abatement - The removal of non-navigation data noise should be accomplished, but only where it is reasonably certain that only non-navigation noise is being purged.

Data Catalog: The PCFR would be the logical place to "inventory" flight recordings. There are two requirements:

- Data Completeness - The existence of any loss of data during the flight should be recorded and output from the PFCR.
- Data Totals - The statistics of data packages in terms of size, granularity, precision, and type should be recorded and output.

Data Security: Only one trivial software requirement is necessary for security purposes, and that is to have more than one copy of the PFCR output available, as a protection against fire or other damages to the master.

3.1.2.3 Duplicate Navigation Data (DND) - The maintenance of a central master file is the best method for sustaining the integrity of the total navigation package. Given that the central file is maintained as n distinct packages of data (perhaps n separate types), the only software requirement for universal distribution purposes is a simple duplication process (the DND). In order to have an additional capability for copying any specified subset of data imaginable, one would require the DND to have an extract logic capability. After thorough consideration is given to the organization of the master file, when creating the specification, it will probably be found that the versatility of an additional extract logic routine will not be required.

3.1.2.4 Create Flight Navigation Logs (CFNL) - Flight summary reports will be required in order that the various navigation design and analysis interests might expedite any required remedial design activity for subsequent flights. This remedial activity must be begun as quickly as possible. The Create Flight Navigation Logs (CFNL) routine (along with the EAR and MQLR) should provide to the analysts, within 24 hours of flight, an overview of flight navigation performance.

The input to the CFNL is all of the data streams within the master file; the output is a set of summary navigation reports presenting various perspectives of thos data streams. To accomplish this output, the following are software requirements;

Navigation Event Catalog: The flight navigation scenario is given by the "navigation state" versus time or versus flight events. The "navigation state" is the set of measured and computed parameters that best characterizes the navigation system. At the minimum, it will be all instrument recordings (which have been formatted, calibrated, and synchronized) and the navigation velocity ($\underline{V}$)/position ($\underline{R}$) state vector. In addition, it will include a characteristic set of intermediate navigation computations, and any computed guidance velocity/position state vectors, (i.e., required velocity/position).

The assumption at this time is that any specific group of analysts will be interested only in a selected subset of the navigation state, which always includes the velocity/position state. Also each group would have its own time frame-of-interest. Therefore, the software requirement of the CNFL is to create the navigation state vs. time and/or vs. flight events stream, and to partition that package of data according to the type of reader; where at least one of the reports would be a summary of the whole navigation state vs. time and/or vs. event.

The events that will dictate navigation state measurements are:

- all mode switches,
- all failures,
- all Redundancy/Management switches,
- any unexpected event, and
- a specified set of time intervals.

For all types of events there need not be a complete navigation state measurement. For example, a Redundancy Management switch might require the reporting of only the instruments involved.

Data Statistical Calculations: The CFNL should also be required to characterize quantitatively each component of the navigation state. A minimum set of calculations for such a characterization includes:

- Specification Check - Each critical data stream is compared with its specification, and deviations are reported.
- Parameters Statistics - The meaningful statistics of any parameter are the statistics of the difference between that parameter and a set of standard references, including

INTERMETRICS INCORPORATED • 701 CONCORD AVENUE • CAMBRIDGE, MASSACHUSETTS 02138 • (617) 661-1840

- spec limits
- expected values
- another parameter that measures the same dimension
- null

The CFNL should have the capability for creating the distribution of any of these relative values.

- Recording Data Statistics - from PFCR
- Noise Statistics - any clearly identified noise should be reported.

The statistical calculations will include, additionally, a set of specialized data statistics as dictated by analysis requirements, and as outlined in the post-flight data processing specification.

3.1.2.5 Event Analysis Routine (EAR) - Certain analyses will be required only when triggered by unusual events. For example, when a particular data stream is suspect, like an accelerometer output, this would dictate the comparison of that stream with other measurers of the same parameter, like the redundant equivalents of the suspect instrument. Because of the complexity of the shuttle instrument package, the number of such causal analyses can be quite large. The number of types of software analyses will, however, be somewhat smaller. The post-flight navigation data preprocessing specification will dictate the specific set of such software analysis capabilities. Following is the set of requirements that should be expected in that specification:

Data Plotting: When an unusual event is found in any data string, the problem is to find the cause or reasons for that event. The obvious recourse is to "plot" the string in which the event occurred against other strings that might give insight. (To "plot", in this case, is to create data pairs that can be presented in graph form or in any manner.) The software requirement is to be able to plot any navigation state parameter against any other state parameter. Examples of the reference parameters are: the velocity/position state; the redundant-equivalent string of data; and the ground-track equivalent string.

Plotting can also include a finer grain of the analysis triggering parameter against time or against events.

Data Analysis: Any data string or the map of two or more data strings can be characterized by certain statistical parameters. Examples constituting the minimum software requirement for the EAR are:

- correlation coefficients,
- means,
- standard deviations, and
- curve fits (e.g., least square).

Complementary Data Reporting: There exists within the master file a complete set of navigation data for all flights. To aid in the diagnosis of unusual events, that file can be used to create a "Diagnosis Book" for each class of unusual events. Included in that book would be such things as:

- calibration data from this flight and all previous flights
- instrument repair history,
- failure history,
- configuration history,

and any other data deemed as a necessary input to diagnosis.

3.1.2.6 Modified Quick-Look Routine (MQLR) - The Modified Quick-Look Routine (MQLR) anticipates that thorough foresight at the time of specification writing is not possible. There will always be a need for quick analyses that are not covered in the preceding QLP routines. What the specification can cover, however, is the scope of MQLR activity, including, for example, the types of analyses that will be accomplished; level of manpower/machine power available; etc. It is expected that the MQLR software capabilities will be variations on existing capabilities like the EAR routines. It is also expected that MQLR requests will expedite revision action against the post-flight navigation data processing specification.

3.1.3 QLP Summary

The quick-look processor, as described in the previous sections, enables a rapid and orderly distribution of information concerning flight navigation performance. Significant cost savings will be the result of implementing the QLP as described because of: the minimization of waiting on the part of design and analysis activities; the elimination of duplicate efforts; and the reduction of post-flight testing to include only that which is absolutely essential.

INTERMETRICS INCORPORATED • 701 CONCORD AVENUE • CAMBRIDGE, MASSACHUSETTS 02138 • (617) 661-1840

3.2 Post-Flight Navigation Processor (PFNP)

The post-flight navigation processor provides the capability of executing the redundancy management and navigation functions subsequent to flight. A general description of this processor and its uses is contained in the following paragraphs.

3.2.1 Functional Requirements

Three types of inputs are needed for the post-flight navigation processor. These include: preset navigation parameters and constants, time history of shuttle navigation sensor data recorded throughout the flight, and ground tracking trajectory data. Specifics as to the data types are given in Section 2.3.

The input data is processed by a post-flight version of the operational redundancy management (RM) and navigation functions. The baseline version of the PFNP should contain an exact duplication of the on-board RM and navigation software (filters, sensor selection logic, navigation mechanization, etc.). The baseline PFNP should be able to reproduce the navigation portion of the actual flight using the recorded data.

The output of the PFNP is a time history of the position, velocity, and attitude of the shuttle and filter-determined variances of the errors in these quantities. In addition, measurement statistics for each sensor can be compiled. The trajectory output of the PFNP is compared with the trajectory prepared by ground tracking equipment to achieve a measure of the performance of the operational software.

Naturally, the post-flight reproduction of the actual test by the baseline processor does not in itself seem too important. If the operational software functions properly, such data would be available from the test. The real value of the PFNP lies in the fact that many alternatives can be exercised using real data collected in an operational environment without repeating the actual test. Hence, the major uses for the PFNP are:

<u>Navigation Algorithm Improvement Studies</u> - Changes in the filters or other navigation algorithms can be made, and the impact of these changes can be tested in the operational environment. Comparison of the altered PFNP output with the trajectory determined by ground tracking equipment can be used as a performance index.

<u>Redundancy Management Testing</u> - It is not expected that the ALT will exercise all aspects of the redundancy management function. By altering the test data, these capabilities can be tested in an operational environment. Also, alternative RM algorithms can be tested. This can include changes to the failure detection and identification (FDI) algorithms.

<u>Subsystem and Software Error Detection</u> - Post-flight processing of the form described in this section is useful for the identification of subsystem and software errors. Three ways in which such errors may be identified are: by observing systematic deviation of PFNP and ground tracking trajectory solutions, by using

expanded state filters, and by comparing sensor measurements not used in the PFNP solution with this solution. The latter is performed by executing the PFNP using all but one set of sensor data. This use of a post-flight processor has been shown to be both practical and successful with the CIRIS system*.

The advantages of developing a post-flight processor for the shuttle ALT test are obvious from its potential uses as described above. The disadvantage might be the development and operational cost associated with this technique. However, this cost is minimal with respect to the cost of the test program or even one test.

Compared to pure simulation, post-flight processing as described in this section has the advantage of using operationally-obtained real data. Modification of a simulator to perform this function is often not possible since the post-flight processor timing must be governed by the recorded data timing instead of by the simulator software. Some of these differences will become more apparent in the software description section.

3.2.2 Software Requirements

The suggested design of the shuttle post-flight navigation processor consists of three separately-operable software modules. These modules are required to perform the following functions:

---

*Radio transponder survey errors, doppler radar anomalous performance, and INS mechanization equation errors were all found using a CIRIS post-flight processor.

a. flight data tape editor (FDTE)

b. parameter file tape generator (PFTG)

c. navigation function processor (NFP)

(1) NFP control

(2) job control card processing

(3) redundancy management and navigation processing

(4) ground tracking comparison module

The system configuration is shown in Figure 3-3. The following subparagraphs specify the required operational functions. It should be noted that much of the text that follows is a function of the method of generating the formatted navigation data. In general, multiple tape input will be considered.

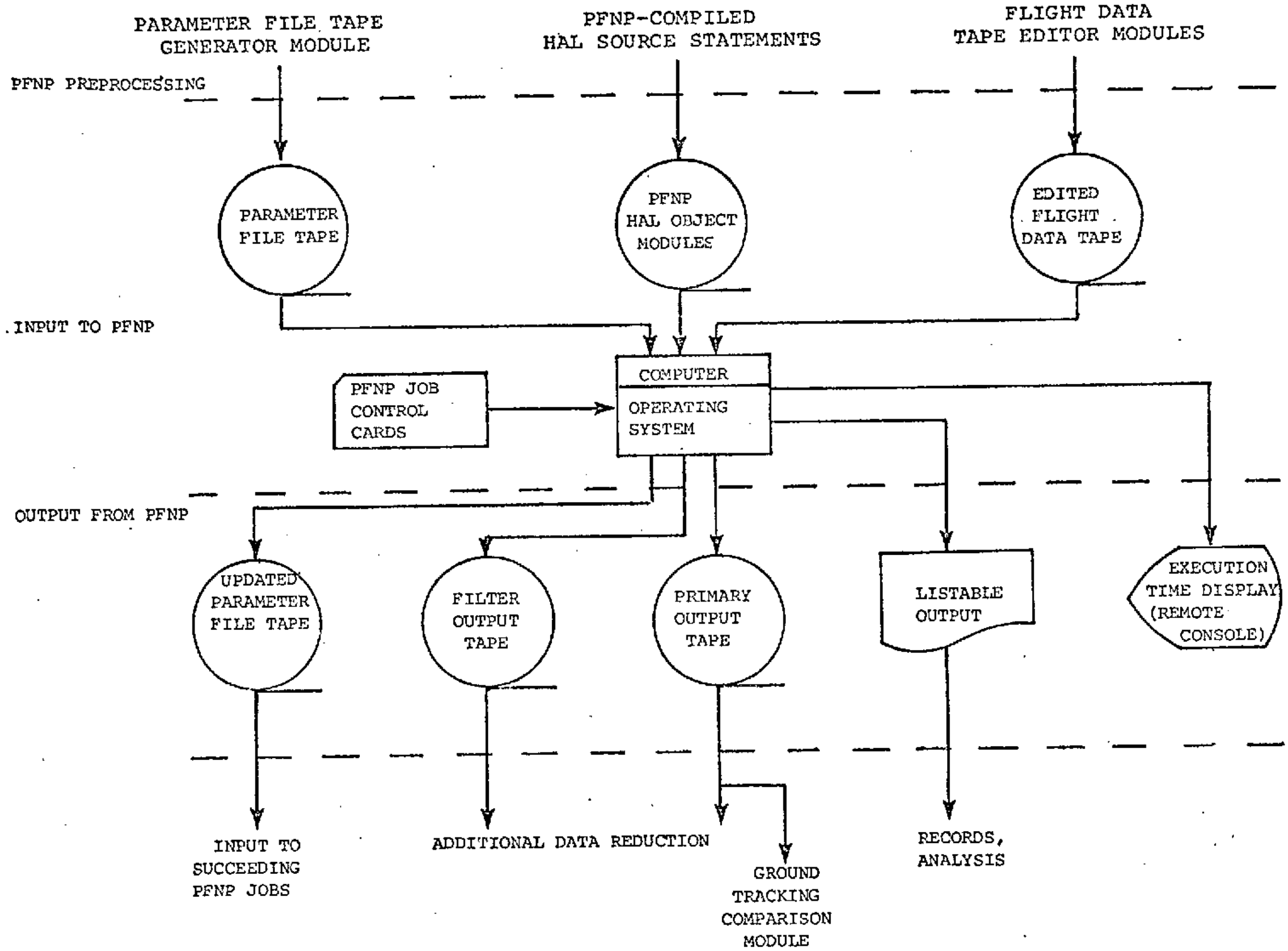


Figure 3-3 Post-Flight Navigation Processor System Configuration

3.2.2.1 Flight Data Tape Editor (FDTE) - This preprocessor module processes the formatted navigation data tapes and creates a new edited flight data tape. This consists of processing the flight data to a state of completeness by simulating failed sensors and deleting data anomalies. The flight data tape editor module has been designed as a separate program for several reasons. In many situations, many different PFNP runs can be made while the data editing needs to be done only once. Also, in general, the same edited data tape can be used for both the PFNP and error isolation processor (see Section 3.3). In addition, the flight data tape editor provides a standard interface for use in generating test tapes used in program checkout.

3.2.2.1.1 Source and Type of Inputs - The inputs to the flight data tape editor are a subset of the formatted navigation data tapes described in Section 3.1.3. Those tapes will have been specifically designed for the purposes of the post-test data processor immediately following the flight. The master of these tapes will be kept in a central file, eliminating a need for rigorous control during post-flight navigation processing.

3.2.2.1.2 Destination and Type of Outputs - The output of the FDTE should consist of error messages and an edited flight data tape or tapes (EFDT). Output data is generated once per execution.

Listable output of messages and processed data should be included as an option.

3.2.2.1.3 Information Processing - A simple conceptual flow diagram of the flight data tape editor is presented in Figure 3-4. The job control functions cards are the first processed to produce a tape identifier. Error messages for the function cards should be included in this processing. Processing is terminated if function card errors are found. Next, the navigation data is read a record at a time. Errors in the records are identified and listable output messages provided. Such errors might include parity errors, sync errors, record length errors, and word or bit pattern errors. If no errors are found, the record is processed. The edited data is then output to a tape and listed, if desired. The next record is then read. At the last tape record, a check for multiple tapes is made, and, if more than one tape is available, a new tape unit is assigned, and processing continues.

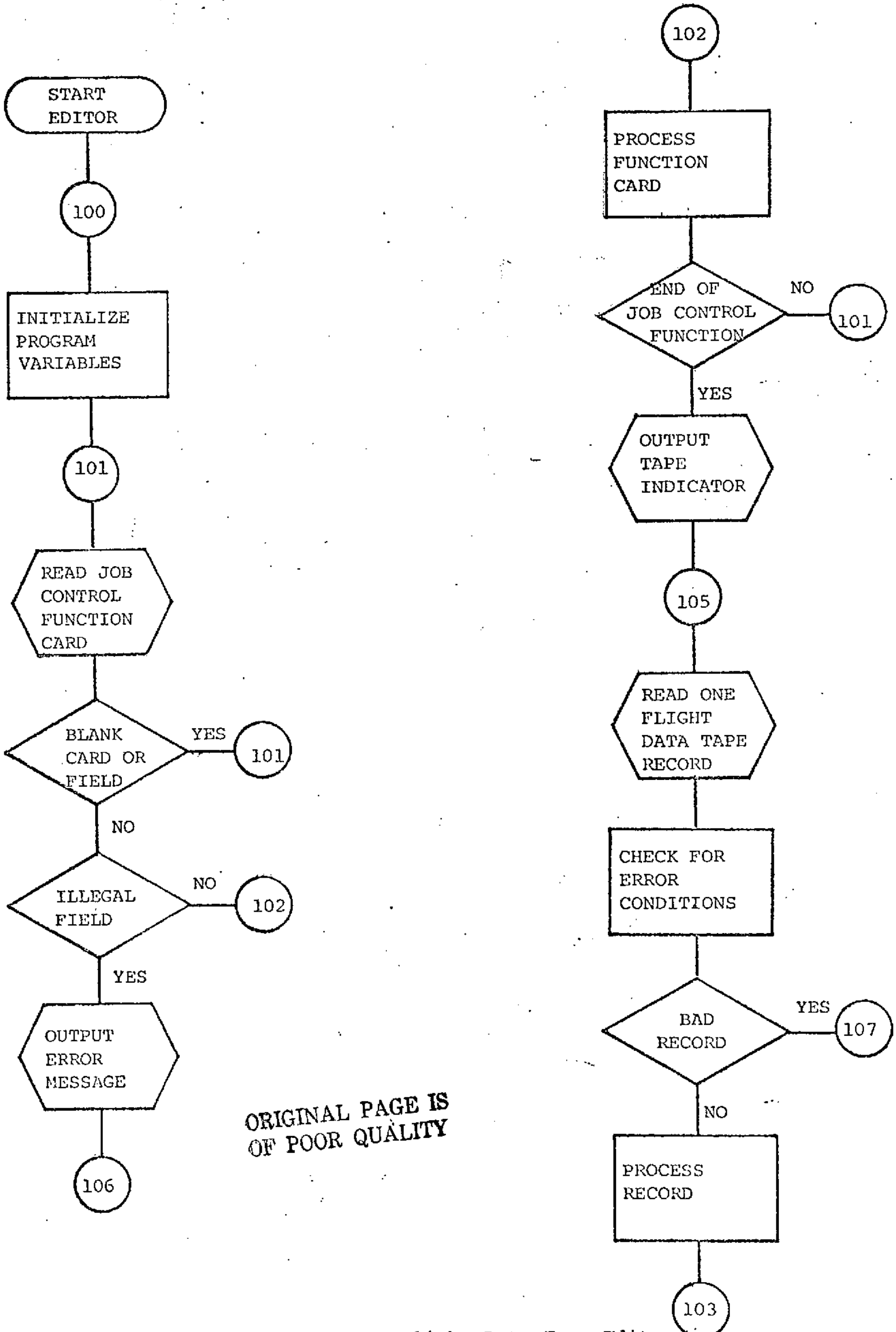


Figure 3-4 Conceptual Flow Diagram for Flight Data Tape Editor

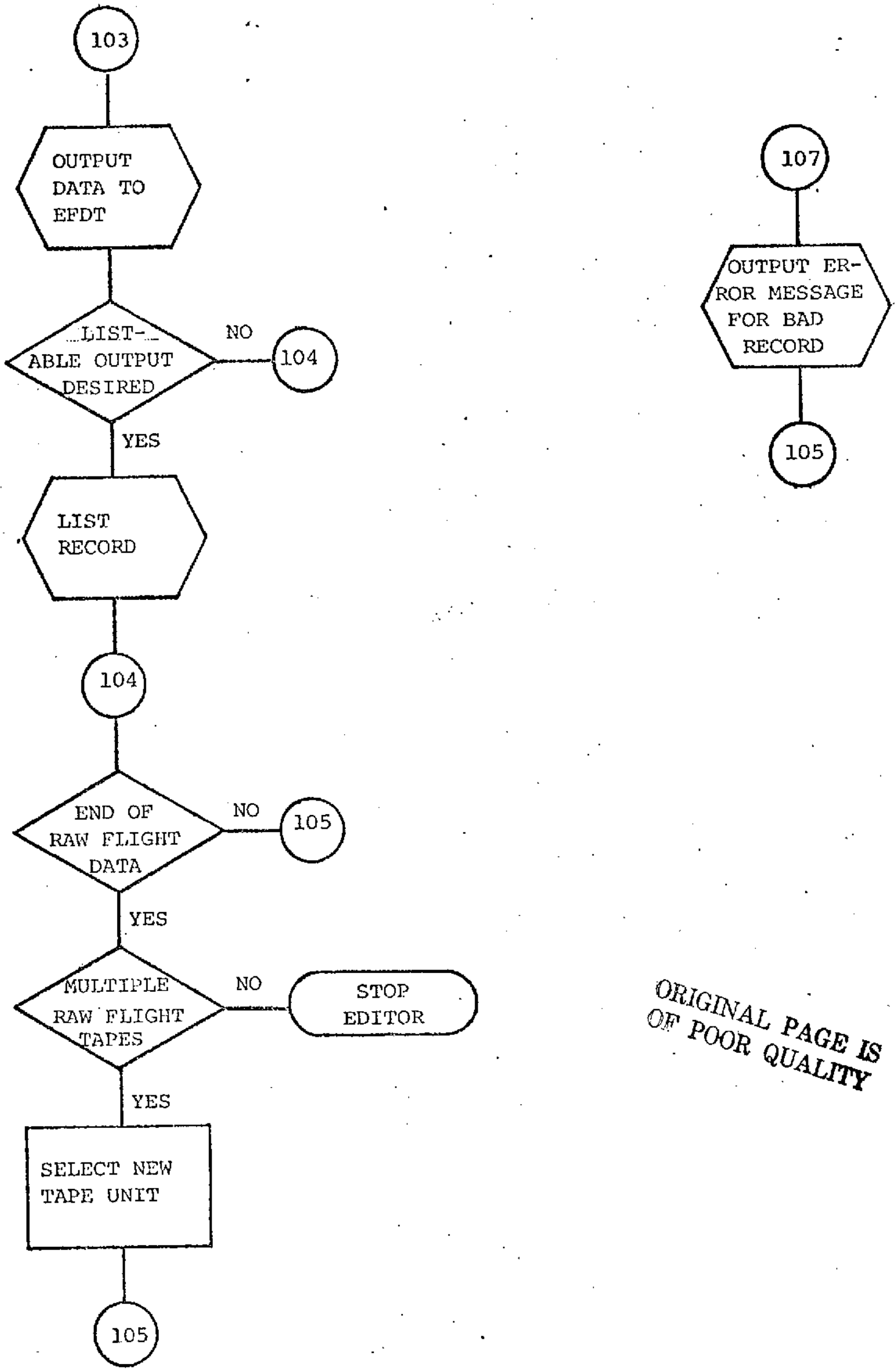


Figure 3-4 (Continued)

3.2.2.2 Parameter File Tape Generator - This preprocessor module creates the fixed data base for the post-flight navigation processor. The data base contains navigation parameter and constants as specified in Section 2.3. The sources of data may be punched cards, paper tape, etc. The parameter file tape generator (PFTG) has been designed as a separate program since a common tape input provides a more convenient method of data base input to the PFNP. As will be seen in Section 3.2.2.3, the inclusion of data base alteration options in the navigation function processor (NFP) allows simple data base changes without the requirement for re-execution of the PFTG.

3.2.2.2.1 Source and Type of Inputs - The input to the parameter file tape generator (PFTG) is of the following three record types:

- job control functions
- data
- job termination

The job control functions perform housekeeping tasks and ultimately result in a tape identity record. This identifier should contain such items as tape generation identifier (PFTG), tape identifier, date, originator identification, etc.

The data records provide for the numeric values that make up the items in the data base (see Table 2-1). The record design should provide the name of the value and the value itself. This syntax allows both ease of data input and verification. No data record order is required since the envisioned design contains a sort routine as part of the baseline. This capability allows easy data input and reduces possible errors.

The job termination records cause the PFTG processor to stop reading records and to generate the parameter file tape (PFT).

3.2.2.2.2 Destination and Type of Output - The outputs from the PFTG module consist of a listing of the contents of the PFT, the tape itself, possible errors in the job control functions, data variable name errors, and data values which are missing.

3.2.2.2.3 Information Processing - A conceptual flow diagram of the parameter file tape generator is shown in Figure 3-5. The job function records are processed in the same manner as in Section 3.2.2.1.3 to produce a tape identifier label. Data records are processed by comparing the inputted variable identifier to a pre-stored table. If a match is found, the inputted value is stored in the appropriate location. During record processing, a count is kept of errors encountered in reading the records. Presence of any errors will inhibit the generation of the parameter file tape; an error summary would be sent to the listable output device, and the job would terminate.

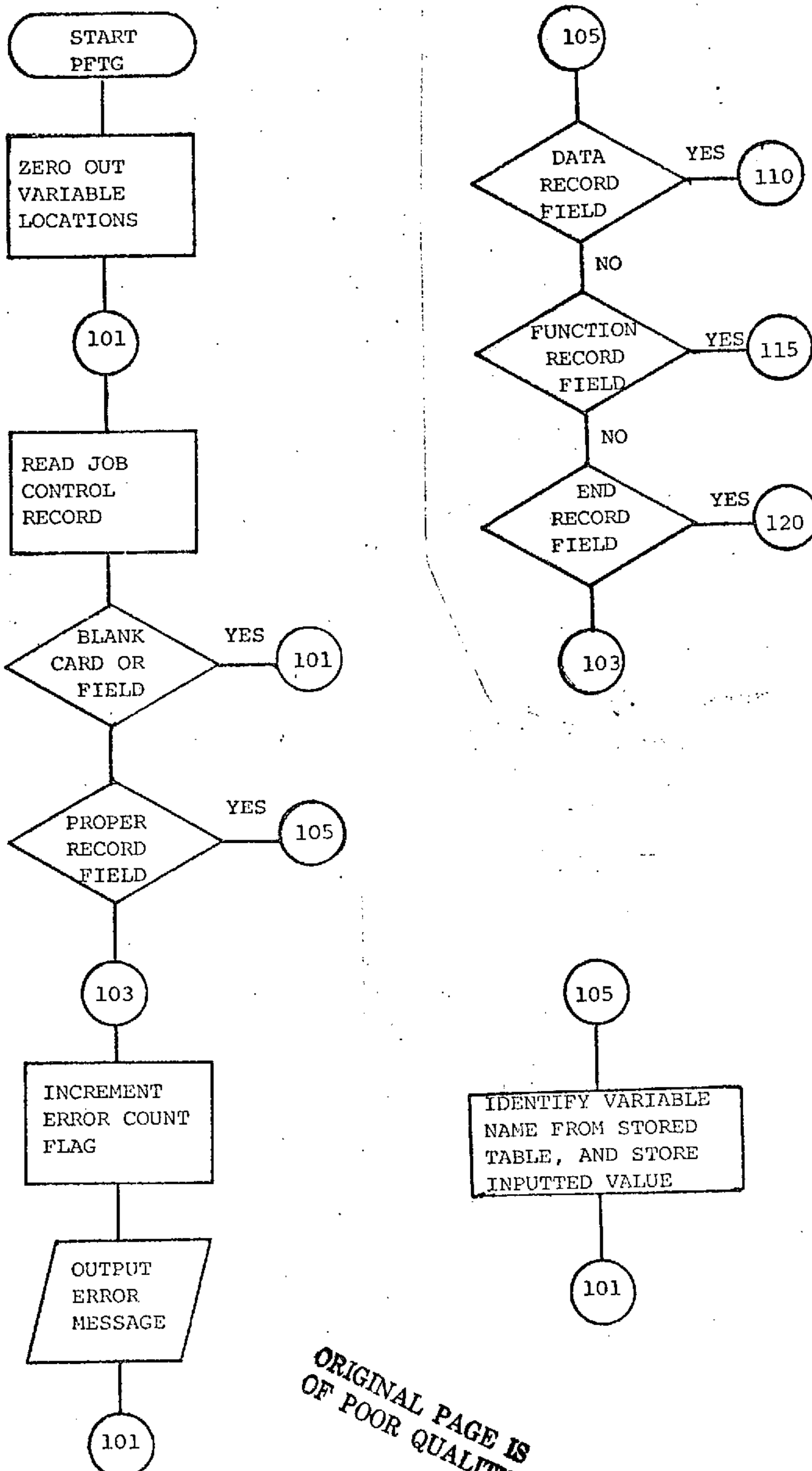


Figure 3-5 Parameter File Tape Generator

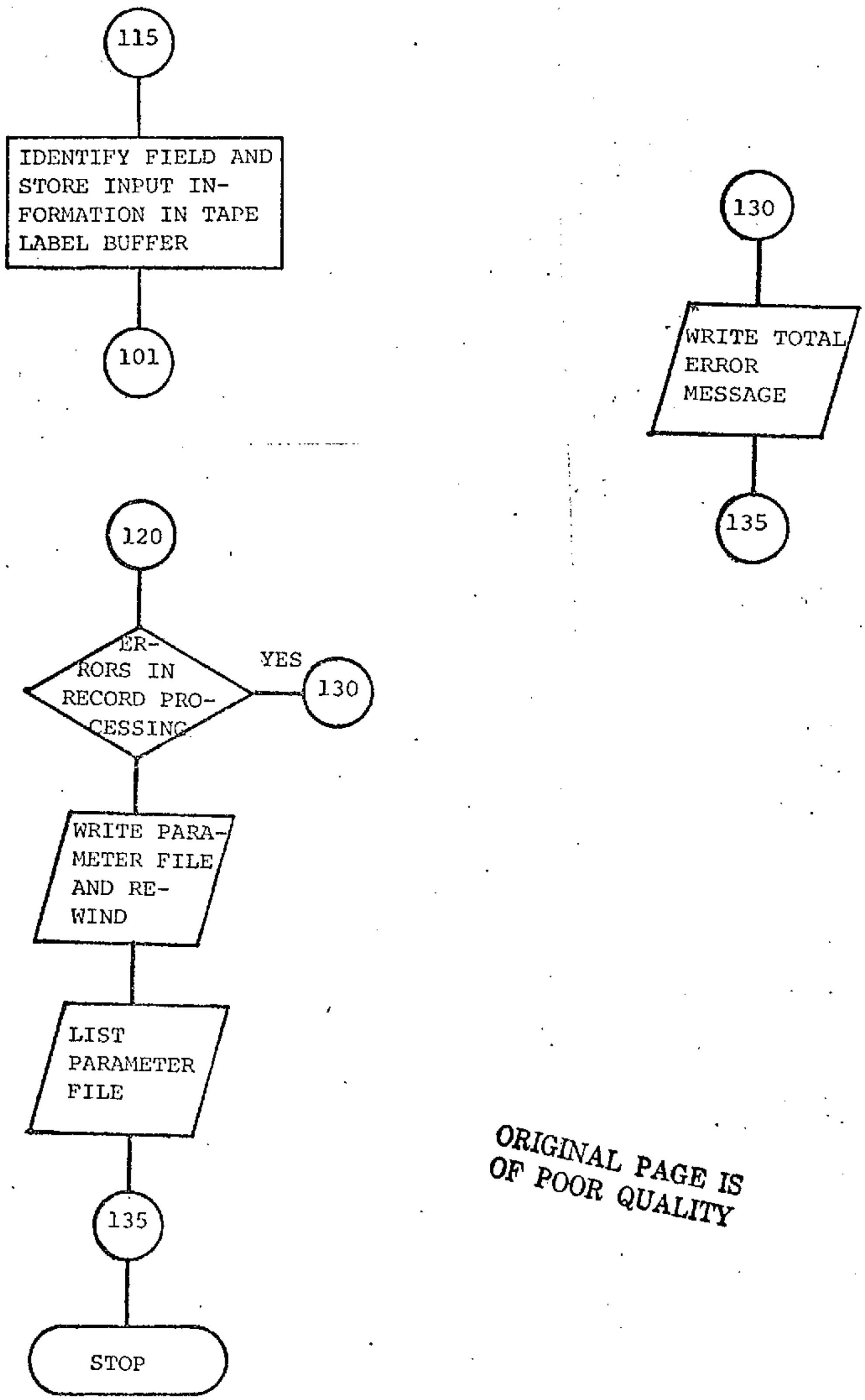


Figure 3-5 (Continued)

If no errors are encountered, a label is written on the tape, followed by the prestored table of variable identifiers and the table of their corresponding values as modified by the data input cards, and the tape is rewound. The contents of the parameter file are then listed for record-keeping purposes, and the job terminates.

3.2.2.3 Navigation Function Processor (NFP) - This main module of the PFNP recreates the execution of the onboard redundancy management and navigation software by operating on the flight recorded data. The parameter base is made available through a parameter file tape (PFT); the contents of this parameter file are modifiable under user control. Processing is directed through user-generated job control cards. In order to offer maximum usefulness, the NFP provides a large degree of flexibility through user options like selection of processing start and stop times, selection of sensor data to be incorporated by the filter, frequency of sensor data processing, etc.

3.2.2.3.1 Source and Type of Inputs - Inputs to the NFP consist of an edited flight data tape, a parameter file tape, and a job control deck. The EFDT is generated by the RFDTE pre-processor (see Section 3.2.2.1). The PFT is generated by the PFTG pre-processor (see Section 3.2.2.2). The job control deck can contain cards of the following types:

- function
- data alteration

- user option generation
- listable output control
- termination

The function card performs administrative duties, leading to identification labels to be placed in any tapes and/or listable output to be generated during the run.

The data alteration cards allow modification of the data base values stored in the parameter file without requiring its regeneration by the PFTG program.

The user option generation card implements user processing directions such as sensor data to be processed, selection output tape generation, etc.

The listable output control card is used to specify the frequency and type of data to be listed during the run.

The termination card indicates the end of the job control deck.

3.2.2.3.2 Source and Type of Outputs - NFP outputs consist of data tapes for further processing or for automatic comparison with outside references, and listable output for quick-look analysis.

All output tapes are labeled with identifiers provided through the job control functions, in order to facilitate correlation of tape data to the conditions of the run in which it was generated. Examples of candidate data for tape output include the filter estimate of shuttle position and velocity for comparison to cinetheodolite and/or C-band radar reference trajectories, and to the onboard computed estimates for verification of compatibility of the operational and post-flight software and evaluation of the effect of limited computer word precision on the performance of the onboard estimator. The Kalman filter state vector and covariance matrix allow insight into the filter performance by observation of the developed cross-correlations between state elements and by comparison of the state variances to the actual level of the filter estimation error as generated from ground tracking reference trajectories. An important index of filter performance is the degree of consistency between the measurement residuals (difference between measured and estimated value of the measurement) and the filter-computed 1σ error in the estimated measurements. Similarly, means and standard deviations can be computed from measurement residuals to evaluate the approximation of the actual residuals to the zero mean random residual measurement error assumed by the filter and the appropriateness of the selected 1σ error level for the measurement. These processed residuals give an indication of the adequacy of the error model being used by the filter.

Listable output would consist of reproduction of the job control card deck, listing of the operating data base, error messages and diagnostics, and a subset of the tape output data to allow quick-look analysis of the NFP performance before embarking on further data reduction.

3.2.2.3.3 Information Processing - A simplified flow diagram for the navigation function processor is presented in Figure 3-6. Following an initialization step, in which administrative functions like I/O unit assignments, etc., are accomplished, the parameter file tape is read. The program then reads the job control cards, identifying their types: function cards are processed to set up label information for output; data cards are read to modify contents of the parameter file data base; generate cards are read to set flags which implement user options to control processing; listable output cards are read to determine run pertinent requirements. Reading of a termination card completes processing of the job control deck. If errors have been encountered in processing the cards, the run terminates following printout of an error summary.

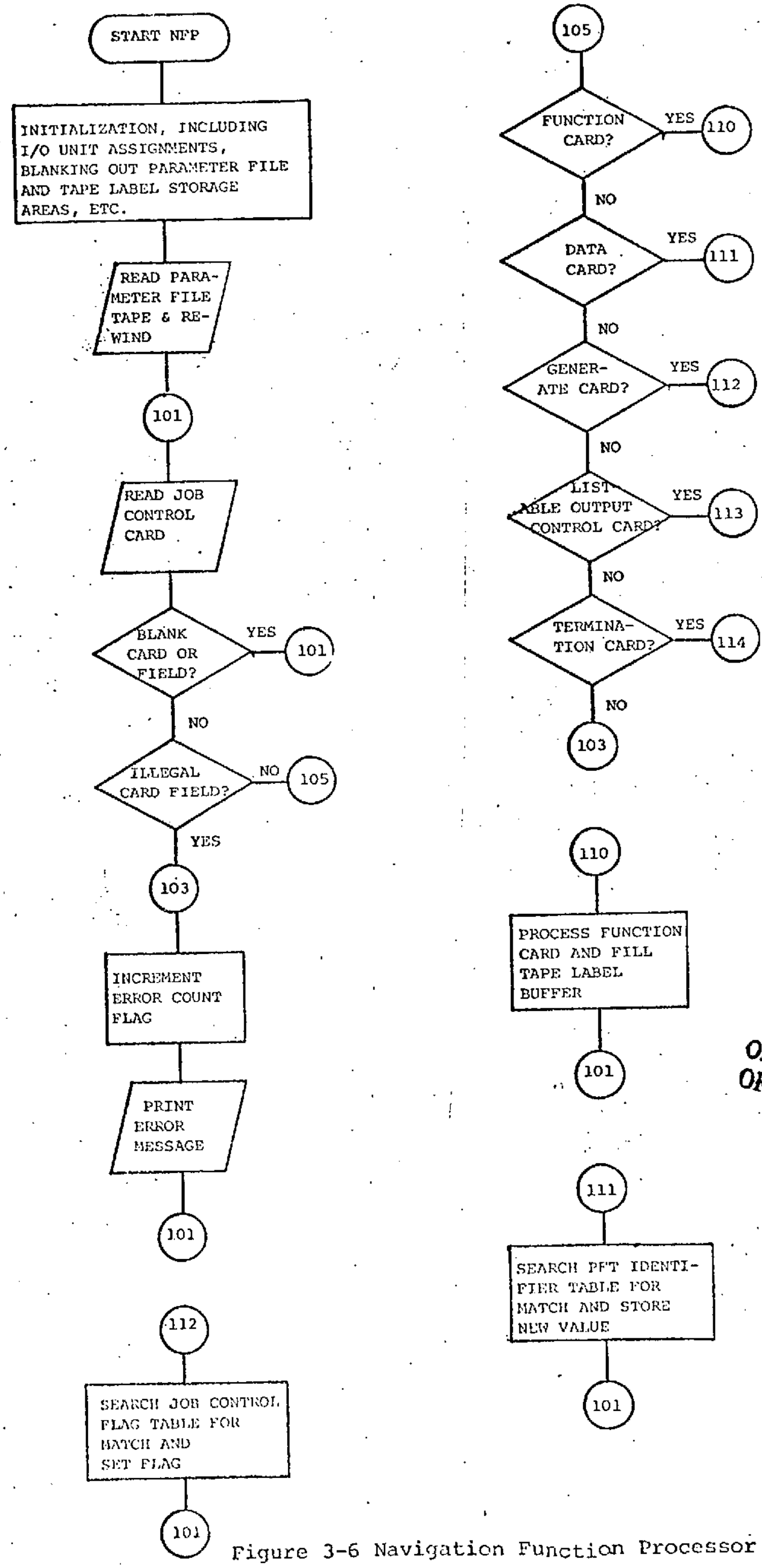


Figure 3-6 Navigation Function Processor

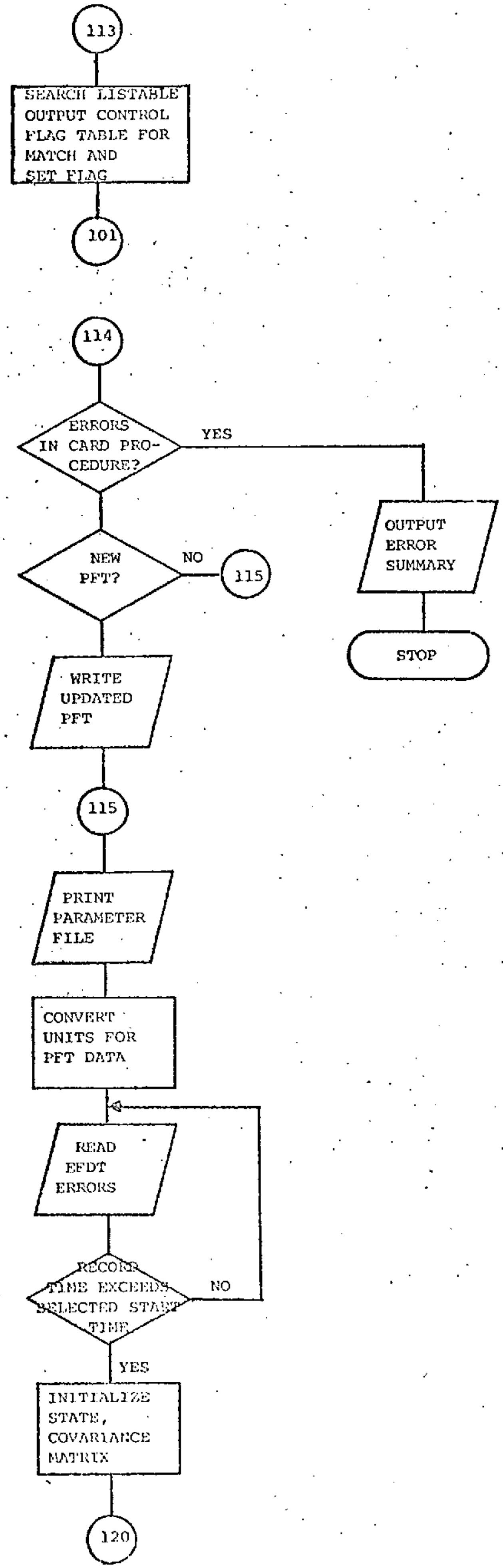
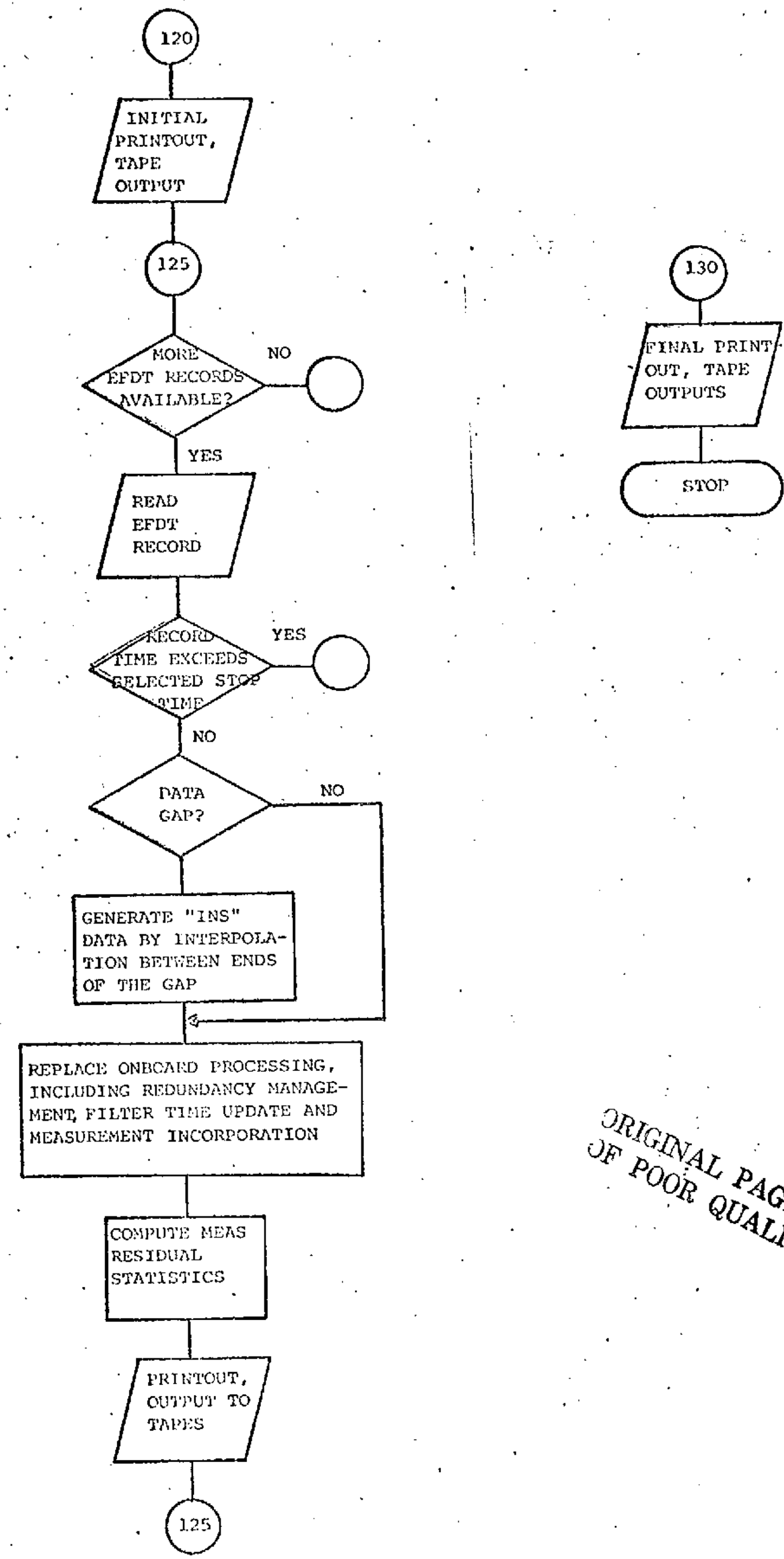


Figure 3-6 (Continued)



ORIGINAL PAGE IS OF POOR QUALITY

Figure 3-6 (Continued)

If so selected, an updated PFT can be generated using the data base modifications input through data alteration cards. The data base is then listed for record-keeping. Parameter units are then converted to a consistent set (°/hr to rads/sec, for example). The program then searches the EFDT for the selected processing start time by reading records until the record time exceeds the desired start time. The Kalman filter state vector and covariance matrix are then initialized and output to tape and the listable output device.

If an end-of-file has not been reached on the EFDT, and the desired stop time has not been exceeded, the program settles into its main loop. The next EFDT record is read. If a data gap exceeding the filter time step constraint is encountered, an interpolated data set is generated for the filter's state and covariance matrix update. The data is then processed by a replica of the on board-redundancy management and navigation software. Output is then made to tape and printer, and the program recycles through its main loop.

On encountering an end-of-file on the EFDT, or on exceeding the desired processing stop time, a final output is made, and the program terminates.

3.2.3 PFNP Summary

The concept and formulation of the post-flight navigation processor described in the previous sections provide not only a post-flight navigation data reconstruction capability, but also a means of altering and testing with flight data the redundancy management and navigation functions. These capabilities offer a significant means of maximizing test data returns while minimizing test operational costs and risk. If formulated using the guidelines of Section 2, modification and checkout of the operational software based on post-test analysis can also be simplified.

3.3 Error Isolation Processor (EIP)

The error isolation processor (EIP) serves a twofold purpose: on the one hand, by estimating the operating error sources, it provides a capability to evaluate the performance of the redundancy management logic; on the other hand, it provides the capability of identifying those error source, or groups of error sources, whose contribution to the system error is critical to the performance of the shuttle navigation function. A general description of this processor and its applications is given below.

3.3.1 Functional Requirements

Two input data types are required for the error isolation processor: modeling parameters, and time history of shuttle navigation sensor data recorded throughout the flight. In addition, ground tracking trajectory data can be used to verify the performance of the processor, or as an external reference source in the estimation process.

The input data is processed by a high-order Kalman filter, modeling all of the significant navigation sensor error sources. In order to make optimal use of all the available information, the filter model could include the significant error sources of each of the three shuttle IMUs. Given a conservative estimate of at least 40 significant error sources for each IMU, the required state dimension, with its related computer storage and computation time requirements, rapidly becomes unmanageable. Furthermore, the errors from the different IMUs can be expected to be highly correlated, given that they are driven by the same environment (gravity errors) and that they maintain a fixed relative orientation. Therefore, their contribution to improved estimation of each other's error sources will be small. It therefore seems more cost-effective to optimally blend each IMU's outputs with the external reference data one at a time. The same processor can still be used by providing for user selection of the IMU data to be processed. The state dimension is reduced to a manageable level (~50 states) with corresponding savings in program run time. Note, however, that a source of information has been neglected (i.e., the other two IMUs) and that, therefore, the processor performance is somewhat degraded versus the theoretically optimal achievable estimation accuracy.

The output of the error isolation processor is a time history of the estimates of vehicle position velocity and attitude and of the modeled nav sensor error sources, and the filter-computed covariance of the estimation errors. The trajectory output can then be compared to the ground tracking equipment trajectory to provide confidence in the filter performance.

The desired output from the processor are the estimates of the operating navigation sensor error sources. Two major uses for this data are presented below:

a. Redundancy Management Software Evaluation - Determination of the IMU error sources provides the basis for RM evaluation by establishing whether a failure had indeed occurred when the software so indicated, or conversely, whether the software failed to react in the face of an actual component failure. On this basis, the RM design can be validated or appropriate modifications justified.

b. Identification of Error Sources Critical to Navigation Performance - An error source, or a linear combination of error sources, is observable to the extent that is contributes to the measured error. Thus, in the event that the navigation filter estimates are not consistent with external measurements, resulting in degraded filter performance, the error isolation processor provides a ready tool for the identification of error sources, modeling errors, etc., causing the poor performance. With its high-order state, and therefore greater error dissemination capability, it allows accurate determination of the error source characteristics and permits clear-cut testing of competing hypothesis.

This processor also offers the possibility of using smoothing techniques to achieve improved estimation accuracy. Smoothing (in the sense of Fracier) represents the combination at any point in time of optimal filters processing the data forward and backward in time. It thus provides, at any point, the optimal estimate derived from all the available data, with significant improvement of the estimation of time-varying errors (for constant errors, the final forward filter estimate represents the result of processing all the available data, and is, therefore, the optimal estimate). The benefits to be provided by smoothing depend then on the extent to which the shuttle nav sensor error sources are expected to be time-varying. In general, most IMU error sources are considered to be constants, and smoothing would not then provide any improved estimation accuracy.

3.3.2 Software Requirements

The suggested design of the shuttle error isolation processor consists of three separately operable software modules:

a. flight data tape editor (FDTE),

b. parameter file tape generator (PFTG),

c. error analysis processor (EAP).

The first two modules follow the same design approach as those described for the PFNP; in fact, a generalized version of these modules could be developed to serve as interfaces for both main processor (PFNP and EIP). Thus, descriptions of the FDTE and PFTG will not be repeated here.

3.3.2.1 Error Analysis Processor - The EAP attempts to estimate the significant error sources of each of the shuttle nav sensors. The EAP follows the same design philosophy as the NFP. It reads the nav sensor data off an edited flight data tape. The EAP parameter base is made available through a parameter file tape; the contents of the file can be modified by the EAP under user control. EAP processing is directed through user-selected job control cards, controlling such options as processing start and stop time, IMU-to-be-processed selection, reference system selection, etc.

3.3.2.1.1 Source and Type of Inputs - Inputs to the EAP consist of the edited flight data tape (EFDT), the parameter file tape (PFT) and a job control deck. The EFDT is generated by the flight data tape editor pre-processor. The PFT is generated by the parameter file tape generator pre-processor. The job control deck serves to specify user options and contains the following types of cards:

- function cards, specifying tape label and listable output page header information
- data alteration cards, specifying modifications to the program data base contained in the parameter file
- generation cards, specifying user-selected processing options
- listable output control cards, specifying the type and frequency of program listable output
- termination card, indicating end of the job control deck

3.3.2.1.2 Source and Type of Outputs - EAP outputs consist of tapes for further processing and listable output for visual analysis.

All tapes are labeled with identifiers provided through the job control function cards. Tape output data consists of the state estimates and covariance matrix at each filter update.

Listable output consists of a subset of the tape output data, provided at lower frequency. Such data might consist of position and velocity estimates and their computed variances, plus measurement residuals, in order to confirm filter convergence. In addition, error messages and diagnostics are provided as required.

3.3.2.1.3 Error Analysis Processing - A simplified flow diagram for the error analysis processor (EAP) is presented in Figure 3-7. The parameter file tape (PFT) is first read to set up the initial program data base. The processing of job control cards, which

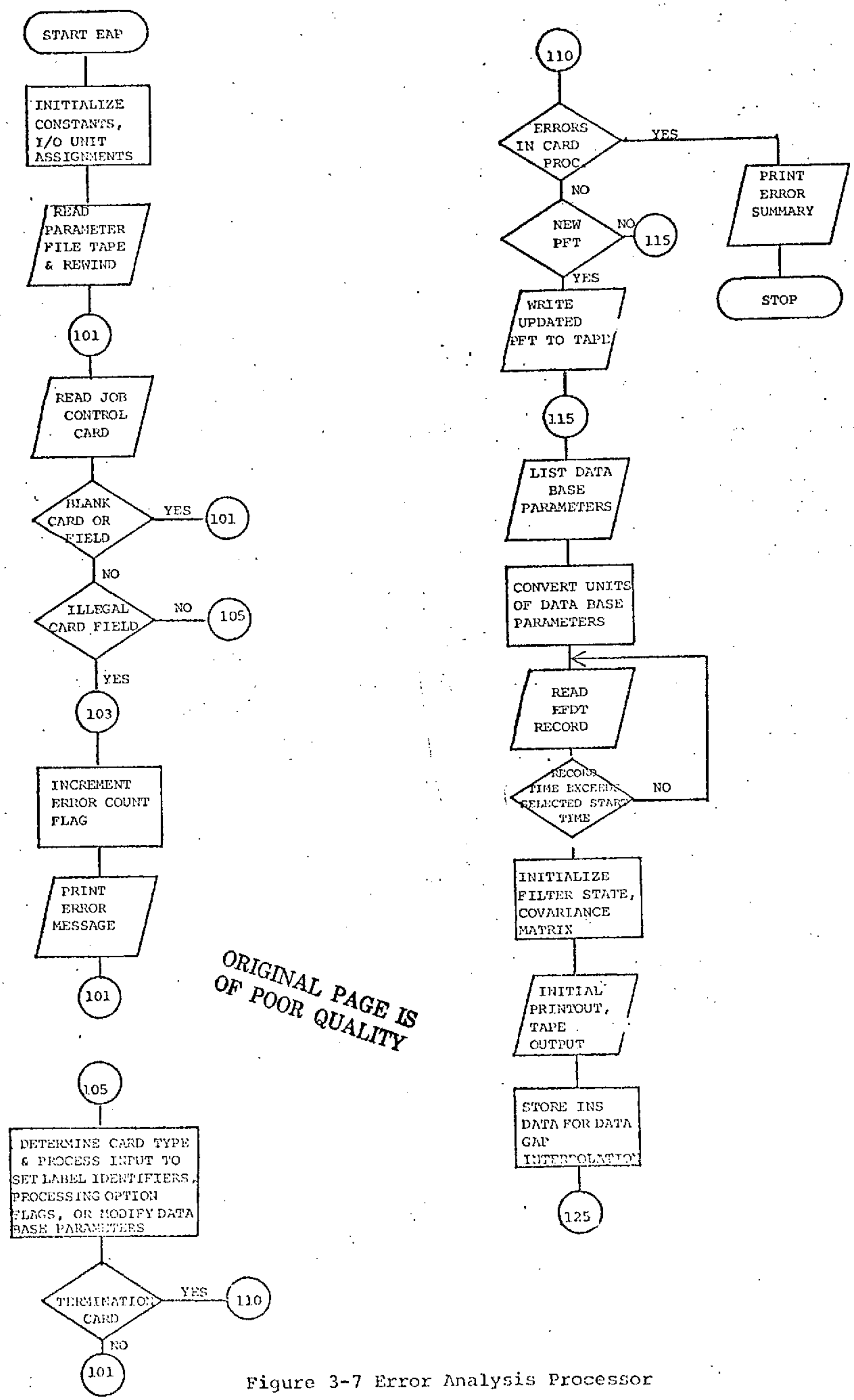


Figure 3-7 Error Analysis Processor

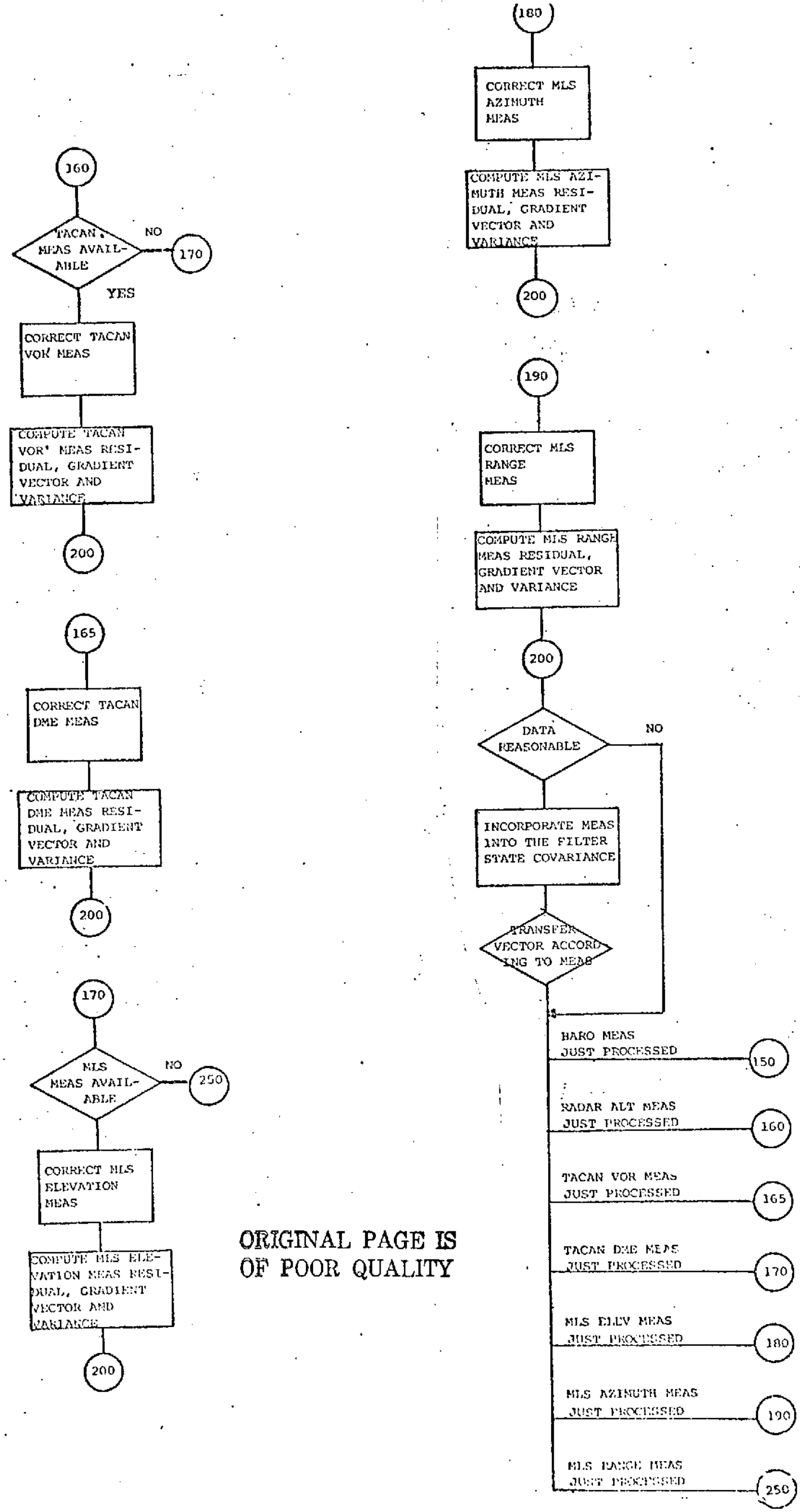


Figure 3-7 (Continued)

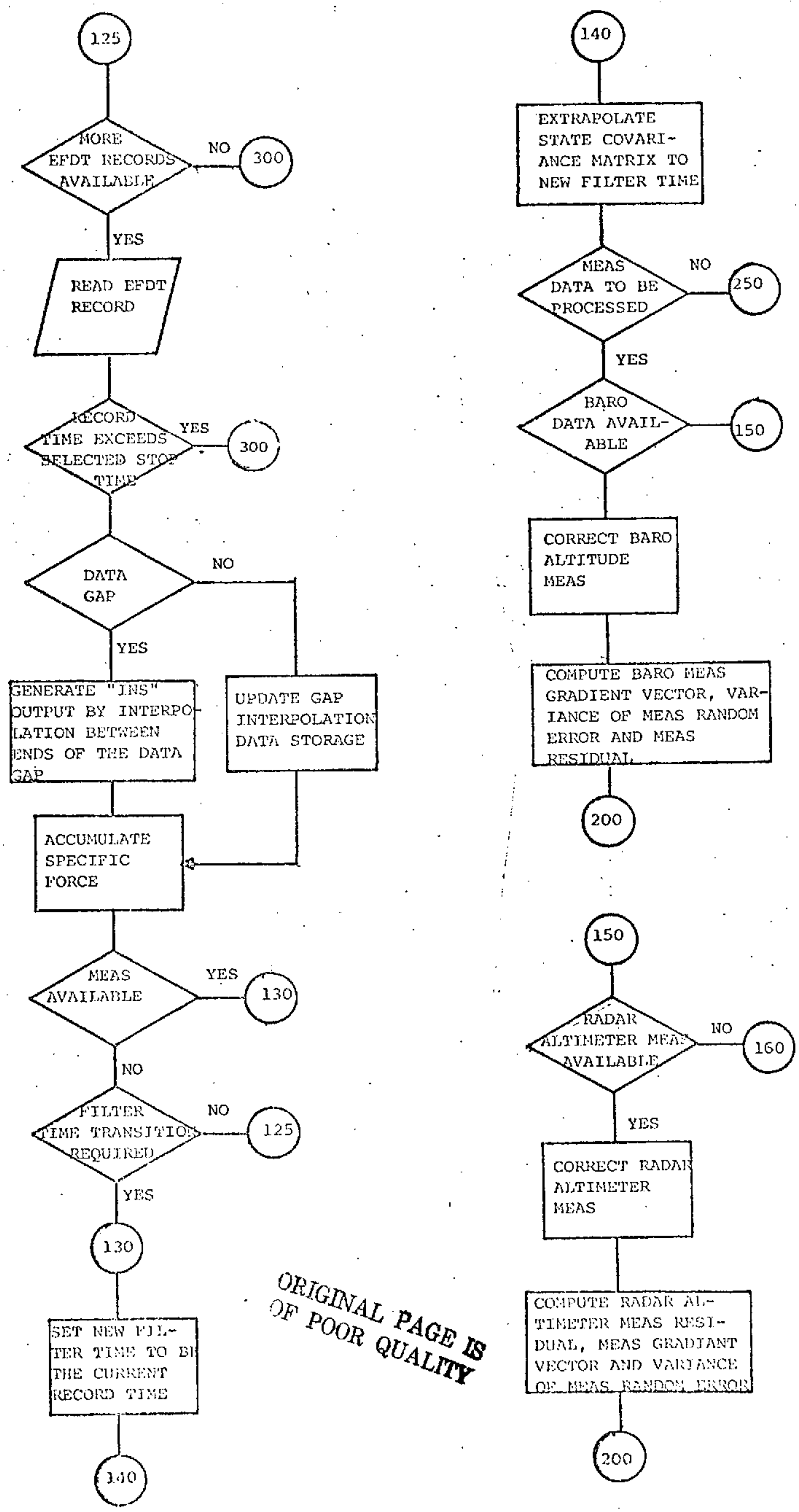


Figure 3-7 (Continued)

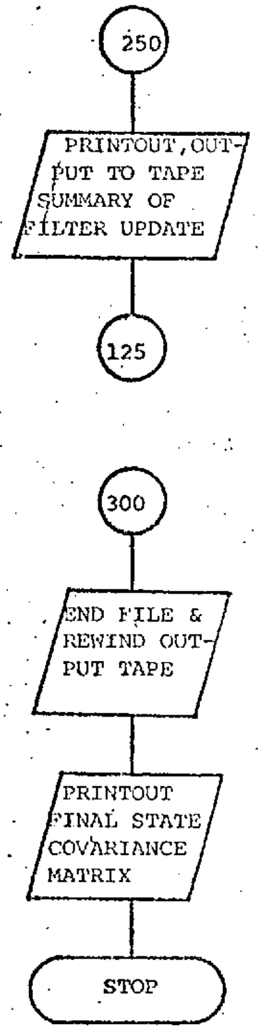


Figure 3-7 (Continued)

ORIGINAL PAGE IS
OF POOR QUALITY

generate header labels, set processing control flags and modify values in the data base, then proceeds similarly as for the NFP (see Section 3.2.2). If errors have been encountered in processing the job control card deck, the program terminates.

If a new PFT is desired, the contents of the parameter file, as modified by the input data alteration control cards, are output to the specified tape. The data base parameters are then listed, following which they are converted from their colloquial type units used for I/O, to a consistent internal processing set (°/hr to rads/sec, $\widehat{\text{min}}$ to rads, g's to ft/sec^2, etc.).

The EFDT is then read until the desired processing start time is found. The filter state and covariance matrix are then initialized. Also to be initialized is a set of INS data to be used as the basis for interpolating across any gaps which may be present in the recorded data.

The program then enters its main loop. If more data is available in the EFDT, the next record is read. If the selected processing stop time has been reached, the program terminates.

The filter is to compute the state transition matrix through the approximation $\phi = I + F\Delta t$, where F is the system dynamics matrix, and Δt is the filter time step. In order to validly neglect the higher-order terms in the approximation, the filter time step must be constrained. Thus, if a data gap is encountered which exceeds the filter time step limit, a set of "pseudo-INS" data must be generated by interpolation for use in the filter time transition. This interpolation would be of higher accuracy than the linear extrapolation used for the state transition matrix computation. The data to be used in this interpolation routine is kept current by updating it each time a record is read.

Some elements of the state transition matrix are functions of integrals of specific force or integrals of products of specific forces (i.e., elements relating gyro g^2 sensitive drifts to platform tilts). Given the rapidly varying character of the specific force vector for the shuttle, the allowable filter time step would be severely constrained. In order to avoid this, the specific force-dependent terms are accumulated at the rate at which the data becomes available, thus still constituting a faithful representation of the shuttle IMU specific force history when they are used at the more leisurely filter update rate. In this fashion, processing time is saved considerably.

Next, if a measurement is available, or if the filter time step constraint is about to be violated, the filter state and covariance matrix are extrapolated to the new filter time.

The order in which the filter is to process the measurement sensor data must be established. One such ordering, organized roughly from the least to the most accurate measurement type, is presented in the flow chart: barometric altimeter altitude, radar altimeter altitude, TACAN VOR meas, TACAN range meas, MLS elevation, MLS azimuth and MLS range. Such an ordering tends to process last the three measurements which are most sensitive to non-linear effects, thus using for them the best available state estimates. Actual processing of the measurements is, of course, dependent on whether the sensor data is available at the selected new filter time.

For each sensor, a measurement correction and calibration step is first executed. This might include compensating the barometric altitude for non-standard temperature effects if temperature readings are available, correcting the radar altimeter, TACAN and MLS measurements for the fact that the sensor antenna is not co-located with the IMU, compensating elevation angle and range measurements for atmospheric refraction effects, etc. Following this, the appropriate measurement residual, measurement gradient vector and variance of the assumed measurement random error is computed.

In order to avoid matrix inversions, the sensor data are treated as sequential scalar measurements. The following processing is common then to all measurements and is called for each measurement. The filter estimated measurement variance ($\underline{h}^T P \underline{h} + r$, where $\underline{h}$ = measurement gradient vector, P = covariance matrix, and r = variance of measurement random error) is computed and used to apply a data-reasonableness test to the measurement residual. If the square of the measurement residual exceeds some multiple of the estimated measurement variance, the measurement is neglected. Such a level might be 16 variances (or 4σ) which contains over 99% of the possible samples for a gaussian distributed random error. If the filter has been properly formulated, the covariance matrix should contain an actual representation of the 1σ level of the error in the estimates, and the test should therefore be able to edit out bad measurements.

If the data validity test is passed, the measurement is used to update the filter state and covariance matrix. Depending on the measurement that has just been processed, and consistent with the pre-established processing order, the next measurement is selected, and the loop repeats until all measurements have been exhausted. Appropriate printouts and outputs to tape are made, and the program loops back and reads a new EFDT record.

When all EFDT records have been exhausted, or when the desired processing stop time is reached, the program terminates.

3.3.3 EIP Summary

The concept and formulation of the error isolation processor described in the previous sections provide the capability of determining magnitudes of particular navigation error sources for any one test. This processor is considered to be the least critical: it provides data that is valuable but not mandatory if the test

INTERMETRICS INCORPORATED • 701 CONCORD AVENUE • CAMBRIDGE, MASSACHUSETTS 02138 • (617) 661-1840

goal is only to determine overall navigation system performance. If additional detail is desired on subsystem performance, the EIP provides a method of obtaining this data.

4. POST PROCESSING DEVELOPMENT PLAN

The development of the navigation data post processor should be divided into three phases:

a. Specification Development Phase - where the preliminary design of the processor will be achieved;

b. Implementation Phase - where the design will be completed, and the software built and verified; and

c. Test Phase - where integration testing will be achieved using data from the orbiter ground checkout facility.

The rationale behind this plan structure and its relationship to other shuttle activity is presented below.

Figure 4-1 shows the three phases and their relationships to the standard NASA development scenario and to the significant shuttle program milestones. The standard scenario is presented in the sections of the Figure 4-1 labelled "Development Activity and Documents" and "Review Milestones". The specification development phase requires an input of the navigation system requirements, design, and test requirements (as shown in the diagram). With those inputs, it will be possible to achieve preliminary design review (PDR) of the post processor within six months. The design

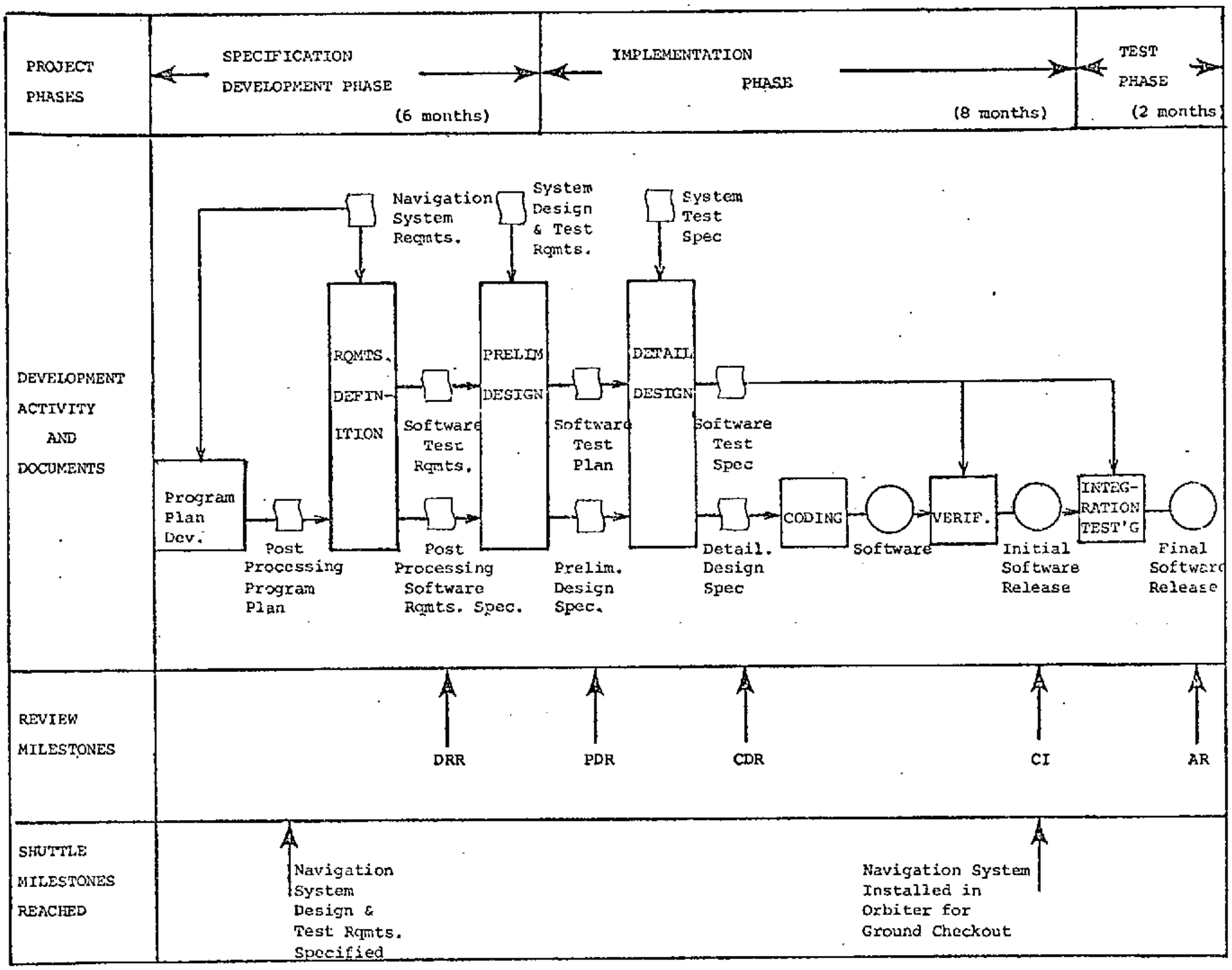


Figure 4-1 Navigation Data Post Processing Development Plan

ORIGINAL PAGE IS OF POOR QUALITY

requirements review (DRR) would be achieved between the second and third month, with the program plan developed in the first month. (The writing of the program plan could be taken as a separate project phase, if so desired.)

The major reason for having a specification development phase separate from an implementation phase is the difficulty in being definitive in the scheduling and the costing of the software package implementation without a prior analysis of the navigation specifications. Hence, the program plan will be updated during the specification writing phase for the purposes of the implementation phase.

The reason for a separate test phase (as opposed to including it in the implmentation phase) is the different phase logistics. Integration testing should be accomplished at the operational test facility using data from the actual vehicle ground checkout; whereas the other activities can be achieved at any location.

A preliminary estimate of the level of effort for the development of various combinations of post-test processing tools is presented in Table 4-1. It should be noted that the level of effort required for the development of more than one form of post processor is not the sum of the man-months required for the independent development of individual processors. This occurs due to the similarity of certain modules in the various post processors.

Table 4-1

Level of Effort (Man-Months) for Combinations of Processor Development

PROCESSOR(S)	SPECIFICATION DEVELOPMENT	IMPLEMENTATION PHASE	TEST PHASE	TOTAL
QLP	6	32	4	42
PFNP	6	32	4	42
EIP	6	32	4	42
QLP & PFNP	11	56	6	73
QLP, PFNP & EIP	16	80	6	102

There are two shuttle program forcing elements that act on the post processing development plan (see the bottom section of Figure 4-1). The primary forcing element is the orbiter vehicle ground checkout. At a point in time when navigation equipment has been installed, the post processor should be ready to demonstrate its capability. According to the schedule in Figure 4-1, 14 months prior to that installation should be the post processing project initiation date. But project initiation requires input of a set of navigation specifications -- the second forcing element. Orbiter ground checkout begins in March, 1976, with the first captive flight scheduled for February, 1977. The navigation system should, therefore, be installed by the summer (say, July) of 1976. This means that May, 1975, would be an appropriate starting date for the post processing development activity, with PDR occurring around November, 1975.

5. SUMMARY AND CONCLUSIONS

The development of navigation data post-flight processors can be, if properly planned and implemented, a significant cost-saving device for the overall shuttle program. Using flight data from the initial test flights, it is possible to accomplish significant design tradeoff analyses of other flights through inexpensive simulation. Those simulations will have the added advantage of using actual operational data. The cost savings of simulation over actual flights are obtained, while the major argument against simulation is answered through the use of actual flight data.

To summarize, the post-flight processing project outlined in this report achieves the following capabilities:

- Rapid reporting of flight performance
- Central control, synchronization, and standardization of data
- Evaluation of alternative navigation schemes and modes
- Performance accuracy analysis

with the following cost and performance advantages to the overall program:

- Early identification of flight anomalies, hence a rapid initiation of remedial design allowing for tighter schedules between flights

- Optimization of flight navigation configuration

  - minimum software requirements
  - minimum calibration and alignment requirements
  - minimum hardware requirements

- Checkout of software without complex flight tests

- Timely response to data requests from analysts, designers, and managers

- Tight control and catalog of data inventory

  - minimum post-flight processing activity
  - no loss or damage of data

- Standard programming language for all analyses

- Common data base for all users

The anticipated development schedule for this project (see Section 4) should be begun by May, 1975, in order to meet the overall schedule objectives of the shuttle program.

INTERMETRICS INCORPORATED • 701 CONCORD AVENUE • CAMBRIDGE, MASSACHUSETTS 02138 • (617) 661-1840

$$
\begin{aligned}
m\ddot{x} &= k^2 \sum_{j=0}^{n-1} mm_j \frac{x_j - x}{\Delta_j^3} + mF_{1x} + F_{2x};\\
m\ddot{y} &= k^2 \sum_{j=0}^{n-1} mm_j \frac{y_j - y}{\Delta_j^3} + mF_{1y} + F_{2y};\\
m\ddot{z} &= k^2 \sum_{j=0}^{n-1} mm_j \frac{z_j - z}{\Delta_j^3} + mF_{1z} + F_{2z}.
\end{aligned}
$$

These equations can be readily derived from the fundamental equation of dynamics for a free material point written in vector form /13

$$m\overline{W} = \overline{F},$$

where $\overline{W}$ is the acceleration of the point; and $\overline{F}$ is the resultant defined by the following sum:

$$\overline{F} = \overline{f} + m\overline{F}_1 + \overline{F}_2.$$

Here $\overline{f}$ is the resultant force of attraction of the celestial body; $m\overline{F}_1$ is the auxiliary mass force caused by the influence of the noncentrality of the planetary force field. This force must be taken into account for motion near the planet; and $\overline{F}_2$ are the forces that do not have a potential function, for example, the force of air drag, which is taken into account for motion in the atmosphere.

After canceling out m, the equations become

$$
\left.
\begin{aligned}
\ddot{x} &= k^2 \sum_{j=0}^{n-1} m_j \frac{x_j - x}{\Delta_j^3} + F_{1x} + \frac{1}{m} F_{2x};\\
\ddot{y} &= k^2 \sum_{j=0}^{n-1} m_j \frac{y_j - y}{\Delta_j^3} + F_{1y} + \frac{1}{m} F_{2y};\\
\ddot{z} &= k^2 \sum_{j=0}^{n-1} m_j \frac{z_j - z}{\Delta_j^3} + F_{1z} + \frac{1}{m} F_{2z}.
\end{aligned}
\right\} \qquad (1.1)
$$

The system of differential equations (1.1) of sixth order is the most general system of equations of motions of artificial spacecraft, including artificial Earth satellites. For assigned initial conditions of motion ($\dot{x}_0$, $\dot{y}_0$, $\dot{z}_0$), it must be integrated by numerical methods. In a special, but very important case when the point travels in the gravity field of only one planet (j) and the perturbing forces $\overline{F}_1$ and $(1/m)\overline{F}_2$ are not taken into account, the system of equations (1.1) can be integrated in the final form.

Let us transfer the origin of coordinates to the center of attraction, that is, let us jet $x_j = y_j = z_j = 0$ (we can neglect the influence of the mass of the AES on planetary motion), then /14
equations (1.1) take on the form of equations of unperturbed motion:

$$\left.\begin{aligned} \ddot{x}+k^2 m_j \frac{x}{r^3}&=0;\\ \ddot{y}+k^2 m_j \frac{y}{r^3}&=0;\\ \ddot{z}+k^2 m_j \frac{z}{r^3}&=0, \end{aligned}\right\} \tag{1.2}$$

where $r = \sqrt{x^2 + y^2 + z^2}$ is the radius-vector of the point under study relative to the center of attraction.

Let us multiply the first equation by y, and the second by x, and let us subtract one from the other. Then,

$$x\ddot{y} - z\ddot{x} = \frac{d}{dt}\left(x\frac{dy}{dt} - y\frac{dx}{dt}\right) = 0.$$

By integrating this equation, we get

$$x\frac{dy}{dt} - y\frac{dx}{dt} = C_3. \tag{1.3}$$

In similar fashion, from the second and third equations, and from the first and third equations, we get

$$y\frac{dz}{dt} - z\frac{dy}{dt} = C_1; \tag{1.4}$$

$$z\frac{dx}{dt} - x\frac{dz}{dt} = C_2. \tag{1.5}$$

These three first integrals of system (1.2) are the integrals of areas.

If we multiply integrals (1.3)-(1.5) by z, x, and y, respectively, and if we add the results, we get

$$C_1x+C_2y+C_3z=0, \tag{1.5'}$$

that is, the motion of a material point acted on by the central force applied at O occurs in a plane passing through point O. This is physically clear since the side forces with respect to the plane containing the radius-vector r of the material point and its velocity vector $\dot{\bar{r}}$ are absolute. The position of this plane in space is entirely defined by the initial conditions of motion, that is, by the initial coordinates of a point and by the initial velocity of the coordinates with integrals (1.3)-(1.5).

By moving in the plane (1.5'), the material point preserves its sectorial velocity constant. We can easily see this since the left-hand sides of Eqs. (1.3)-(1.5) are the projections of the vector $\bar{r} \times \dot{\bar{r}}$ onto the coordinate axes. According to the definition, the vector product of two vectors is equal in magnitude (or modulus) to the product of their moduli by the sine of the angle between them, that is, equal to the area of the parallelogram constructed on these vectors. /15

In our case, $\bar{r}$ is the radius-vector of the material point and $\dot{\bar{r}}$ is its velocity. The area of a parallelogram constructed on these vectors is numerically equal to double the sector velocity. This means that under the effect of the gravitational force of one center (one planet), the material point will move along a plane curve, preserving its sectorial velocity constant and, therefore, preserving its projections onto the coordinate axes constant.

Let us denote the sectorial velocity, that is, the increment in area A swept out by the radius-vector of a traveling point per unit of time by dA/dt. Let us find the value of the constant 1/2 C, which it is equal to:

$$\begin{aligned}\frac{dA}{dt}&=\frac{1}{2}C=\\&=\frac{1}{2}\sqrt{(y_0\dot{z}_0-z_0\dot{y}_0)^2+(z_0\dot{x}_0-x_0\dot{z}_0)^2+(x_0\dot{y}_0-y_0\dot{x}_0)^2},\end{aligned} \tag{1.6}$$

where

$$C=\sqrt{C_1^2+C_2^2+C_3^2}.$$

After integration, we get

$$A = \frac{1}{2} C (t - t_0),$$

that is, the area of sector A increases in proportion to time t (t_0 is the initial moment of time reference). Obviously, the projection of the area of sector A onto any of the coordinate planes will vary according to the same linear law. The constants C_1, C_2, and C_3 are the projections of the doubled sectorial velocity C onto the coordinates planes Oyz, Oxz, and Ozy. If C_1, C_2, and C_3 are known, then therefore not only the sectorial velocity 1/2 C is known, but also defined is the orientation of the plane in which the point (of the AES) will move. The constants C_1, C_2, and C_3 are usually replaced by the more graphic parameters C, i, and Ω, of which: i is the inclination of the orbital plane to the principal coordinate plane Oxy (if we consider the equatorial coordinate system, then the inclination of the orbital plane to the equatorial plane); and Ω is the angle in the plane Oxy between the axis Os and the line of intersection of the planes of the equator and the orbit (nodal line), where the point at which the satellite travels from bottom upwards (ascending node) is selected for the reference of the angle on the nodal line.

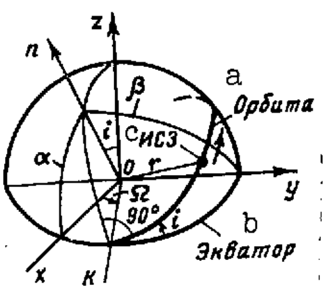


Fig. 1.1. Scheme of motion of AES relative to selected reference system.

Key: a. Orbit; b. Equator; c. AES

From the spherical triangles (see Fig. 1.1) xnK and ynK (the arc nK is equal to 90°) we find /16

$$\cos\alpha = \sin\Omega \sin i;$$
$$\cos\beta = -\cos\Omega \sin i.$$

Therefore, the integrals of areas (1.3)-(1.5) with new constants will be of the form

$$y\dot{z} - z\dot{y} = C \sin\Omega \sin i;$$
$$z\dot{x} - x\dot{z} = -C \cos\Omega \sin i;$$
$$x\dot{y} - y\dot{x} = C \cos i.$$

The longitude Ω of the ascending node will vary within the limits from 0 to 360°, and the orbital inclination i will vary from 0 to 180°. If $0 < i < 90°$ satellite motion will be rectilinear, that is, it will be executed in the same direction in which the longitude increases; if $90° < i < 180°$, the motion will be retrograde.

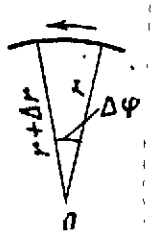


Fig. 1.2. For derivation of law of areas.

The system of equations (1.2) gives us yet another integral -- the integral of energy. In order to obtain this integral, let us multiply the first equation by $\dot{x}$, and the second by $\dot{y}$, and the third by $\dot{z}$, and let us add the results, resulting in the equation

$$\dot{x}\ddot{x}+\dot{y}\ddot{y}+\dot{z}\ddot{z}+\frac{k^2mj}{r^3}(x\dot{x}+y\dot{y}+z\dot{z})=0.$$

Integrating, we arrive at the following equality:

$$\frac{1}{2}(\dot{x}^2+\dot{y}^2+\dot{z}^2)-\frac{k^2mj}{\sqrt{x^2+y^2+z^2}}=H,$$

where H is the constant.

The expression in the parentheses is v^2 -- the velocity squared; therefore, in the final form the integral of energy will be

$$\frac{v^2}{2}-\frac{k^2mj}{r}=H.$$

Along the orbit the sum of the kinetic and potential energies of the AES as it moves in the central field remain constant. By using the integrals of area and energy, we can solve the problem of the motion of the material point acted on by the central force [58].

Let us write out the law of areas (1.6) in the coordinates r and ϕ (angle of rotation of the radius-vector r in the orbital plane):

$$r^2\dot{\varphi}=C. \qquad (1.6')$$

The same result for these orbits is obtained directly from Fig. 1.2, namely /17

$$r(r+\Delta r)\frac{\Delta\varphi}{\Delta t}\approx C=2\dot{A}.$$

As $\Delta t \to 0$, this expression is transformed into (1.6').

Assuming $v^2 = \dot{r}^2 + r^2\dot{\phi}^2$ and $k^2 m_j = \mu$ we can write the equation of energy as:

$$\frac{\dot{r}^2}{2} + \frac{r^2}{2}\dot{\varphi}^2 - \frac{\mu}{r} = H.$$

Canceling out $\dot{\phi}$ and dt by means of Eq. (1.6') we get

$$\frac{C^2}{2r^4}\left(\frac{dr}{d\varphi}\right)^2 + \frac{C^2}{2r^2} - \frac{\mu}{r} = H.$$

Let us introduce the variable $z = C/r$, then

$$\left(\frac{dz}{d\varphi}\right)^2 + z^2 - \frac{2\mu}{C}z = 2H,$$

or

$$\left(\frac{dz}{d\varphi}\right)^2 = 2H + \frac{\mu^2}{C^2} - \left(z - \frac{\mu}{C}\right)^2.$$

Furthermore, we introduce the constant e:

$$2H + \frac{\mu^2}{C^2} = \frac{\mu^2}{C^2}e^2.$$

The constant e enables us to reduce the equation obtained above to the form

$$\frac{dz}{\sqrt{\frac{\mu^2}{C^2}e^2 - \left(z - \frac{\mu}{C}\right)^2}} = \pm d\varphi.$$

After integration we get

$$\varphi + D = \int \frac{dz}{\sqrt{\frac{\mu^2}{C^2}e^2 - \left(z - \frac{\mu}{C}\right)^2}} = \arccos\frac{z - \frac{\mu}{C}}{\frac{\mu}{C}e};$$

$$\cos(\varphi + D) = \frac{z - \frac{\mu}{C}}{\frac{\mu}{C}e}$$

and

$$z - \frac{\mu}{C} = \frac{\mu}{C}e\cos(\varphi + D),$$

where D is the fifth arbitrary constant.

Setting $p = C^2/\mu$ and returning to the variable r, we get /18

and

$$\frac{1}{r} - \frac{1}{p} = \frac{1}{p} e \cos(\varphi + D)$$

$$r = \frac{p}{1 + e\cos(\varphi + D)}. \tag{1.7}$$

This is the equation of a second-order curve in polar coordinates r and ϕ, related to the focus. In it e is the eccentricity and p is the parameter of the curve.

For the ellipse we have

$$p = a_e(1 - e^2),\ e < 1,\ H < 0;$$

for the hyperbola

$$p = a_h(e^2 - 1),\ e > 1,\ H > 0,$$

where a_e is the semi-major axis of the ellipse; and a_h is the actual semi-axis of the hyperbola.

For the parabola $e = 1$ and $H = 0$.

When the angle $\phi = \phi_\pi = -D$, the polar radius r will be at a minimum. This is the perigee of the orbit if the Earth is the center of attraction. When $\phi = \phi_\alpha = -D + \pi$; the polar axis in the case of an ellipse will have a maximum. This is the apogee of the orbit.

The angle $\vartheta = \phi + D = \phi - \phi_\pi$ is called the true anomaly.

It remains to find the dependence of ϑ on time. It will differ for the ellipse, the parabola, and the hyperbola.

Based on the equation

$$r^2 d\varphi = C dt,$$

where

$$r = \frac{p}{1 + e\cos\vartheta},$$

$$C = \sqrt{\mu p}, \tag{1.7'}$$

we get

$$\frac{d\vartheta}{(1+e\cos\vartheta)^2}=\frac{\sqrt{\mu}}{p^{3/2}}\,dt. \tag{1.8}$$

Instead of ϑ we introduce the new variable $\eta = \tan\frac{\vartheta}{2}$; then

$$\vartheta = 2\,\mathrm{arctg}\,\eta;$$
$$d\vartheta=\frac{2d\eta}{1+\eta^2};$$
$$\cos\vartheta=\frac{1-\eta^2}{1+\eta^2}.$$

The left-hand side of Eq. (1.8) can be transformed thusly /19

$$\frac{d\vartheta}{(1+e\cos\vartheta)^2}=\frac{2d\eta}{(1+\eta^2)\left(1+e\frac{1-\eta^2}{1+\eta^2}\right)^2}=\frac{2(1+\eta^2)\,d\eta}{[1+\eta^2+e(1-\eta^2)]^2}=$$
$$=\frac{2(1+\eta^2)d\eta}{[1+e+(1-e)\eta^2]^2}=\frac{2}{(1+e)^2}\cdot\frac{1+\eta^2}{(1+\gamma\eta^2)^2}\,d\eta,$$

where

$$\gamma=\frac{1-e}{1+e}.$$

With reference to this transformation, the initial equation (1.8) takes on the form

$$\frac{1+\eta^2}{(1+\gamma\eta^2)^2}\,d\eta=\frac{\sqrt{\mu}(1+e)^2}{2p^{3/2}}\,dt. \tag{1.9}$$

Depending on the magnitude and sign of γ, the solutions (or integrals) of this equation will differ. The most practical value is found in the case $\gamma \neq 0$, which breaks down into two variants:

$$\gamma<0,\ e>1$$

and

$$\gamma>0,\ e<1.$$

However, it is simplest to integrate Eq. (1.9) when $\gamma = 0$ (parabola). For it, we get

$$(1+\eta^2)d\eta=\frac{2\sqrt{\mu}}{p^{3/2}}dt.$$

Integration yields a formula for calculating the time of flight along a parabola

$$\frac{2\sqrt{\mu}}{p^{3/2}}(t-\tau)=\operatorname{tg}\frac{\vartheta}{2}+\frac{1}{3}\operatorname{tg}^3\frac{\vartheta}{2},$$

where τ is the time of transit of the spacecraft across the orbital perigee (ϑ = 0)

For the ellipse and the hyperbola, the derivation of the dependence of the angle ϑ on time t by means of variable η is both complex and is an artificial process. However, referring to the fundamental nature of the resulting solutions, we present it in full [58]. First of all we note that

$$\frac{1+\eta^2}{(1+\gamma\eta^2)^2}=\frac{\frac{1}{\gamma}(1+\gamma\eta^2)+1-\frac{1}{\gamma}}{(1+\gamma\eta^2)^2}=\frac{1}{\gamma}\cdot\frac{1}{1+\gamma\eta^2}+\frac{\gamma-1}{\gamma}\cdot\frac{1}{(1+\gamma\eta^2)^2},$$

therefore

$$\int\frac{(1+\eta^2)\,d\eta}{(1+\gamma\eta^2)^2}=\frac{1}{\gamma}\int\frac{d\eta}{1+\gamma\eta^2}+\frac{\gamma-1}{\gamma}\int\frac{d\eta}{(1+\gamma\eta^2)^2}. \tag{1.10}$$

Let us apply to the first integral in the right-hand side of the equation the formula of integration by parts /20

$$\int\frac{d\eta}{1+\gamma\eta^2}=\frac{\eta}{1+\gamma\eta^2}+2\int\frac{\gamma\eta^2 d\eta}{(1+\gamma\eta^2)^2}=\frac{\eta}{1+\gamma\eta^2}+$$
$$+2\int\frac{d\eta}{1+\gamma\eta^2}-2\int\frac{d\eta}{(1+\gamma\eta^2)^2},$$

hence

$$\int\frac{d\eta}{(1+\gamma\eta^2)^2}=\frac{1}{2}\cdot\frac{\eta}{1+\gamma\eta^2}+\frac{1}{2}\int\frac{d\eta}{1+\gamma\eta^2}.$$

Then Eq. (1.10) becomes

$$\int\frac{(1+\eta^2)\,d\eta}{(1+\gamma\eta^2)^2}=\frac{\gamma-1}{2\gamma}\cdot\frac{\eta}{1+\gamma\eta^2}+\frac{\gamma+1}{2\gamma}\int\frac{d\eta}{1+\gamma\eta^2}. \tag{1.11}$$

Using (1.9) and (1.11), we can write the integral of Eq. (1.8) as:

$$\frac{\sqrt{\mu}}{p^{3/2}}(t-\tau)=\frac{1}{(1+e)^2}\left\{\frac{\gamma-1}{\gamma}\,\frac{\eta}{1+\gamma\eta^2}+\frac{\gamma+1}{\gamma}\int_0^\eta\frac{d\eta}{1+\gamma\eta^2}\right\}. \tag{1.12}$$

Since

$$\frac{\gamma-1}{\gamma}=\frac{-2e}{1-e}, \qquad \frac{\gamma+1}{\gamma}=\frac{2}{1-e},$$

then (1.12) can be written thusly:

$$\frac{\sqrt{\mu}}{p^{3/2}}(t-\tau)=\frac{2}{(1+e)(1-e^2)}\left\{\int_0^\eta\frac{d\eta}{1+\gamma\eta^2}-e\,\frac{\eta}{1+\gamma\eta^2}\right\}. \tag{1.13}$$

From Eq. (1.13) we directly get the formulas for calculating the time of flight in the case of an elliptical and a hyperbolic orbit.

For the ellipse ($e < 1$, $\gamma > 0$), we have

$$\int\frac{d\eta}{1+\gamma\eta^2}=\frac{1}{\sqrt{\gamma}}\operatorname{arctg}(\eta\sqrt{\gamma}).$$

We introduce the new variable E, using the equality

$$\eta\sqrt{\gamma}=\operatorname{tg}\frac{E}{2}. \tag{1.14}$$

Therefore, we have

$$E=2\operatorname{arctg}(\eta\sqrt{\gamma});$$
$$\frac{2\eta}{1+\gamma\eta^2}=\frac{2}{\sqrt{\gamma}}\sin\frac{E}{2}\cos\frac{E}{2}=\frac{1}{\sqrt{\gamma}}\sin E$$

and Eq. (1.13) becomes /21

$$\frac{\sqrt{\mu}(t-\tau)}{a_e^{3/2}(1-e^2)^{3/2}}=\frac{2}{(1+e)(1-e^2)}\left\{\frac{1}{\sqrt{\gamma}}\operatorname{arctg}\eta\sqrt{\gamma}-\frac{e}{2\sqrt{\gamma}}\sin E\right\}$$

or

$$\frac{\sqrt{\mu}}{a_e^{3/2}}(t-\tau)=E-e\sin E, \qquad (1.13')$$

where E is the eccentric anomaly.

Eq. (1.13') for calculating the time of flight in an elliptical orbit is called Kepler's equation.

The physical meaning of E is clear from Fig. 1.3, in which, besides the trajectory of the satellite motion in an ellipse, is shown the trajectory of motion along a circle of point A that has an x-value that is identical with the satellite:

$$FB=OB-OF \quad \text{or} \quad r\cos\vartheta=a_e\cos E-a_e e.$$

But from the equation of an ellipse it follows that

$$r=\frac{a_e(1-e^2)}{1+e\cos\vartheta}.$$

Dividing the two last equations by each other, we get

$$\cos\vartheta=\frac{(1+e\cos\vartheta)(\cos E-e)}{1-e^2}.$$

Therefore,

$$\cos E=\frac{e+\cos\vartheta}{1+e\cos\vartheta}$$

or

$$\frac{1-\operatorname{tg}^2\frac{E}{2}}{1+\operatorname{tg}^2\frac{E}{2}}=\frac{e+\frac{1-\operatorname{tg}^2\frac{\vartheta}{2}}{1+\operatorname{tg}^2\frac{\vartheta}{2}}}{1+e\frac{1-\operatorname{tg}^2\frac{\vartheta}{2}}{1+\operatorname{tg}^2\frac{\vartheta}{2}}}=$$

$$=\frac{1+e-(1-e)\operatorname{tg}^2\frac{\vartheta}{2}}{1+e+(1-e)\operatorname{tg}^2\frac{\vartheta}{2}}.$$

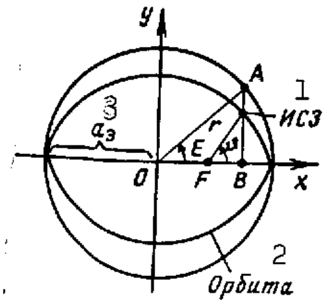


Fig. 1.3. For deriving the formula relating the angles of the eccentric anomaly E and the true anomaly ϑ

Key: 1. AES; 2. Orbit; 3. a_e

After obvious transformations we have /22

$$1+\mathrm{tg}^2\frac{E}{2}+e+e\,\mathrm{tg}^2\frac{E}{2}-\mathrm{tg}^2\frac{\vartheta}{2}+e\,\mathrm{tg}^2\frac{\vartheta}{2}=$$
$$=1-\mathrm{tg}^2\frac{E}{2}+e-e\,\mathrm{tg}^2\frac{E}{2}+tg^2\frac{\vartheta}{2}-e\,\mathrm{tg}^2\frac{\vartheta}{2};$$
$$(1+e)\,\mathrm{tg}^2\frac{E}{2}=(1-e)\,\mathrm{tg}^2\frac{\vartheta}{2}$$

and, finally,

$$\mathrm{tg}\,\frac{E}{2}=\sqrt{\frac{1-e}{1+e}}\,\mathrm{tg}\,\frac{\vartheta}{2}. \qquad (1.14')$$

Thus, we get Eq. (1.14'), which gives us a relationship between E and ϑ.

For the hyperbola (e > 1, $\gamma < 0$), we have

$$\int\frac{d\eta}{1+\gamma\eta^2}=\frac{1}{2\sqrt{-\gamma}}\ln\frac{1+\sqrt{-\gamma}\eta}{1-\sqrt{-\gamma}\eta}.$$

We introduce the variable q, defined by the equality

$$\mathrm{tg}\,\frac{q}{2}=\sqrt{-\gamma}\eta=\sqrt{-\frac{1-e}{1+e}}\,\mathrm{tg}\,\frac{\vartheta}{2}.$$

After obvious transformations of (1.13), we get

$$\frac{\sqrt{\mu}}{a_r^{3/2}}(t-\tau)=e\,\mathrm{tg}\,q-\ln\mathrm{tg}\left(\frac{\pi}{4}+\frac{q}{2}\right),$$

or

$$t-\tau=\frac{a_r^{3/2}}{\sqrt{\mu}}(e\,\mathrm{sh}\,q-q).$$

Thus, the problem of integrating the equations of unperturbed motion (1.2) has been solved. As to be expected, its solution depends on six parameters: a, e, τ, i, Ω, and ϕ_π (or ω), and in rectangular coordinates -- on x_0, y_0, z_0, $\dot{x}_0$, $\dot{y}_0$, and $\dot{z}_0$. The trajectories are in the form of ellipses (periodic solution) or hyperbolas (aperiodic solution). A singular case is the parabola (e = 1) and a particular case of the ellipse is the circle (e = 0).

For the circular orbit we have

$$E=\vartheta=\varphi;\ a=R;\ t=\frac{R^{3/2}}{\sqrt{\mu}}\varphi,$$

where ϕ is the center of the angle.

The period of revolution for elliptical orbits, as follows from Kepler's equation is /23

$$T=\frac{2\pi}{\sqrt{\mu}}a^{3/2}$$

and does not depend on eccentricity.

We note that for elliptical orbits Eq. (1.8) can be readily integrated by means of the substitution

$$\operatorname{tg}\frac{\vartheta}{2}=\sqrt{\frac{1+e}{1-e}}\operatorname{tg}\frac{E}{2}.$$

By differentiating this formula, we get

$$\sec^2\frac{\vartheta}{2}\,d\vartheta=\sqrt{\frac{1+e}{1-e}}\sec^2\frac{E}{2}\,dE$$

or

$$\left(1+\operatorname{tg}^2\frac{\vartheta}{2}\right)d\vartheta=\sqrt{\frac{1+e}{1-e}}\left(1+\operatorname{tg}^2\frac{E}{2}\right)dE.$$

Hence it follows that

$$d\vartheta=\frac{\sqrt{1-e^2}}{1-e\cos E}\,dE.$$

On the other hand, from the equation $\cos E=\frac{e\mp\cos\vartheta}{1+e\cos\vartheta}$ we have

$$\frac{1}{1+e\cos\vartheta}=\frac{1-e\cos E}{1-e^2}.$$

Using this expression, we can define the radius-vector of an elliptical orbit in terms of the eccentric anomaly, that is,

$$r=a(1-e\cos E).$$

Further, from Eq. (1.8), with reference to these formulas, we have

$$\frac{\sqrt{\mu}}{a^{3/2}}\,dt=(1-e\cos E)\,dE.$$

Integration gives us Kepler's equation (1.13').

Calculation Formulas for Unperturbed Elliptical Orbit

Above we obtained the following fundamental formulas, which can be used in calculating elliptical motion:

$r^2\dot{\vartheta}=C$ — integral of areas

$\frac{v^2}{2}-\frac{\mu}{r}=H$ — integral of energy

$r=\frac{p}{1+e\cos\vartheta}$ — general formula for any motion expressing the radius vector of a satellite in terms of p, e and ϑ; /24

$t-\tau=\frac{a^{3/2}}{\sqrt{\mu}}(E-e\sin E)$ — is Kepler's equation for calculating the time of flight in elliptical motion; and

$p=\frac{C^2}{\mu}$ — is the parameter of the orbit

and $e=\sqrt{1+\frac{2HC^2}{\mu^2}}$ is eccentricity.

In practice it often becomes necessary to use a series of formulas deriving from the fundamental relations or obtained from geometrical considerations. We present some of them.

We will assume that the point where the satellite is inserted into orbit has the coordinates: geocentric latitude ψ_0 and longitude λ_0. The azimuth of the absolute satellite velocity at this point will be denoted by δ_0. From these quantities we can calculate the inclination i of the orbit to the equatorial plane and the argument of the latitude at the moment of insertion, determining the distance of point B of the insertion into orbit from the orbital node (Fig. 1.4). This argument (u_0) is expressed by the formula $u_0 = \omega + \vartheta_0$, where ω is the argument of the perigee (π) of the orbit.

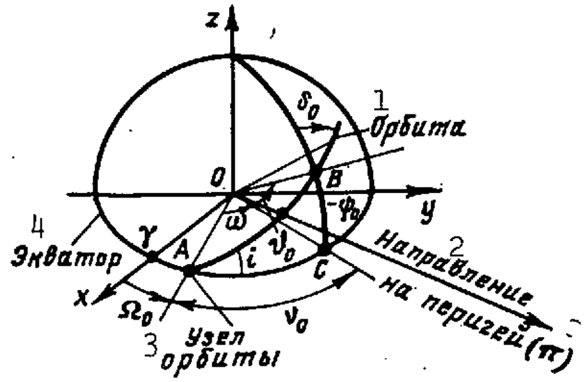


Fig. 1.4. Elements of orbit of an AES.

Key: 1. Orbit; 2. Direction toward perigee (π); 3. Orbital node; 4. Equator

Examining the spherical triangle ABC, we can obtain the following formulas:

$$\cos i = \sin \delta_0 \cos \psi_0; \qquad (1.15)$$

$$\left.\begin{aligned} \sin u_0 &= \frac{\sin \psi_0}{\sin i}; \\ \operatorname{ctg} u_0 &= \frac{\cos \delta_0}{\operatorname{tg} \psi_0}. \end{aligned}\right\} \qquad (1.16)$$

The angular distance of meridian of the point of insertion from the orbital node can be defined from the same triangle: /25

$$\sin \nu_0 = \sin \delta_0 \sin u_0; \qquad (1.17)$$

$$\cos \nu_0 = \frac{\cos u_0}{\cos \psi_0}. \qquad (1.18)$$

We will measure the longitude of the orbital node from the point of vernal equinox. Here it can be determined as the direct ascension of the orbital node. At the moment of insertion into orbit, we have

$$\Omega_0 + \nu_0 = s_0, \qquad (1.19)$$

where s_0 is the sidereal time at the meridian of the point of insertion B into orbit at the moment of insertion; and ν_0 is the arc from the ascending node of the orbit to the meridian of the point B.

Hence we can determine the arc Ω_0. Obviously, in order to obtain the longitude of the node λ measured from some terrestrial meridian, it is necessary to add to Ω_0 the quantity $(-s_0')$, where s_0' is the corresponding sidereal time for the selected meridian.

The resulting formulas (1.15-1.19) are valid, of course, not only for the point of insertion, but for any point on any orbit.

Now let us find expressions for the absolute coordinates x, y, and z, and their time derivatives in terms of the orbital elements Ω, ω, i, a, e, and τ. To do this, let us consider the spherical triangles xAB, yAB and zAB formed by the coordinate axes and the arc AB. Using the theorem of sines for the direction cosines of the radius-vector r of point B, we get

$$\left.\begin{aligned}\cos(rx) &= \cos u \cos \Omega - \sin u \sin \Omega \cos i;\\ \cos(ry) &= \cos u \sin \Omega + \sin u \cos \Omega \cos i;\\ \cos(rz) &= \sin u \sin i.\end{aligned}\right\} \tag{1.20}$$

Therefore,

$$\left.\begin{aligned}x &= r(\cos u \cos \Omega - \sin u \sin \Omega \cos i);\\ y &= r(\cos u \sin \Omega + \sin u \cos \Omega \cos i);\\ z &= r \sin u \sin i.\end{aligned}\right\} \tag{1.21}$$

Differentiating these equations with respect to t, we have

$$\left.\begin{aligned}\dot{x} &= \dot{r}\frac{x}{r} - r(\cos \Omega \sin u + \sin \Omega \cos u \cos i)\dot{\vartheta};\\ \dot{y} &= \dot{r}\frac{y}{r} - r(\sin \Omega \sin u - \cos \Omega \cos u \cos i)\dot{\vartheta};\\ \dot{z} &= r \cos u \sin i \cdot \dot{\vartheta} + \frac{z}{r}\dot{r}.\end{aligned}\right\} \tag{1.22}$$

Here

$$u = \omega + \vartheta; \quad \vartheta = \vartheta(t); \quad r = r(t).$$

The resulting formulas (1.21), (1.22), (1.13'), (1.7'), and and (1.14), that is, the expressions of the absolute coordinates x, y, z, and their derivatives $\dot{x}$, $\dot{y}$, and $\dot{z}$ in terms of their orbital parameters (elements) Ω, ω, i, a, e, τ and time t are the solution to the system of equations of unperturbed motion of a cosmic body (Earth satellite). For each point in space at which the satellite is located, from formulas (1.21) and (1.22), by knowing the five geometric parameters defining the orbit and the one kinematic parameter (τ), for any moment of time we can calculate the absolute coordinates and the satellite velocity. /26

Note that often instead of parameter τ the more convenient parameter M_0 appears.

Let us denote $$M = E - e \sin E = n(t - t_0) + M_0,$$

where

M is the mean anomaly;

$n=\sqrt{\mu}\,a^{-3/2}$ is the mean motion;

t_0 is any specific (initial) instant of time;

$M_0=n(t_0-\tau)$ is the mean anomaly for the instant t_0, the so-called mean anomaly of the epoch; and

M_0 is the parameter replacing τ and characterizing the value of M for the instant t_0.

Let us find the expressions for the orbital parameters in terms of the elements at the end of the powered flight section: v_0, θ_0 (angle of inclination of velocity vector to the horizon), and r_0.

Based on the formulas

$$r=\frac{a(1-e^2)}{1+e\cos\vartheta};\quad r^2\dot{\vartheta}=C;\quad p=\frac{C^2}{\mu}$$

the complements of the velocity vector can be expressed thusly:

in terms of the radius-vector

$$v_r=\dot{r}=\sqrt{\frac{\mu}{p}}\,e\sin\vartheta; \tag{1.23}$$

in terms of the transversal

$$v_n=r\dot{\vartheta}=\sqrt{\frac{\mu}{p}}(1+e\cos\vartheta). \tag{1.24}$$

Hence the modulus of velocity is

$$v^2=\frac{\mu}{p}(1+2e\cos\vartheta+e^2); \tag{1.25}$$

the velocity at the perigee (maximum) is /27

$$v_\pi=\sqrt{\frac{\mu}{p}}(1+e), \tag{1.26}$$

and the velocity at the apogee (minimum) is

$$v_\alpha=\sqrt{\frac{\mu}{p}}(1-e). \tag{1.27}$$

For the radius-vector of a satellite at the perigee (minimum), we have

$$r_\pi=a(1-e); \tag{128}$$

at the apogee, we have

$$r_\alpha=a(1+e). \tag{1.29}$$

Substituting the values of v_π and r_π into the energy equation with reference to the fact that $p = a(1 - e^2)$, we get

$$\frac{v^2}{2}-\frac{\mu}{r}=H=-\frac{\mu}{2a}. \tag{1.30}$$

Thus, total satellite energy depends only on the semi-axis a of the satellite orbit.

Replacing in Eq. (130) v and r by v_0 and r_0 at the point of insertion into orbit, we get

$$a=\frac{r_0}{2-k}, \tag{1.31}$$

where

$$k=\frac{r_0 v_0^2}{\mu}.$$

The quantity k is the square of the ratio of velocity v_0 at the initial point to the velocity in the circular orbit with radius r_0.

Actually, when $k=\frac{r_0 v_0^2}{\mu}=1$, based on Eq. (1.31) $a = r_0$, which corresponds to the circular orbit. In this case,

$$v_0=\sqrt{\frac{\mu}{r_0}}=\sqrt{\frac{\mu}{a}}=\text{const.}$$

The flight velocity v_{ci} [ci = circular] in the circular orbit (r = const) is constant and can be expressed by the formula:

$$v_{ci}=\sqrt{\frac{\mu}{r}}.$$

This velocity is called the circular velocity at the distance r from the center of attraction.

If in Eq. (1.31) we set k = 2, the semi-axis of the ellipse a becomes equal to infinity, that is, the elliptical orbit is transformed into a parabolic orbit. Here we have /28

$$v_{par}=\sqrt{\frac{2\mu}{r}}$$

which is called the parabolic velocity at distance r from the center of attraction.

If we take r = R (where R is the radius of the Earth) in the formula for the circular velocity, we get the first escape velocity

$$v_1=\sqrt{\frac{\mu}{R}};$$

and on analogy

$$v_2=\sqrt{\frac{2\mu}{R}}$$

is called the second escape velocity.

Let us find the expressions for the orbital element p and e in terms of the initial parameters r_0, v_0 and θ_0 defining the ellipse. Since

$$r_\pi v_\pi = \sqrt{\mu p} = rv \cos\theta$$

(where θ is the angle of inclination of the velocity vector to the local horizon), the orbital parameter is

$$p = r_0 k \cos^2\theta_0. \tag{1.32}$$

Replacing in Eq. (1.32) parameter p by its expression $\frac{r_0}{2-k^2}(1-e^2)$, we get the following formula for eccentricity e:

$$e = \sqrt{(k-1)^2 \cos^2\theta_0 + \sin^2\theta_0}. \tag{1.33}$$

The angular distance (true anomaly) of the point of insertion from the orbital perigee will be determined from the polar equation of the ellipse by

$$\cos\vartheta_0 = \frac{1}{e}(k\cos^2\theta_0 - 1), \tag{1.34}$$

and the angular distance of the perigee from the orbital node will be

$$\omega = u_0 - \vartheta_0. \tag{1.35}$$

If at the end of the powered section $\theta_0 = 0$, the orbital perigee coincides with the end of the section, and here $\vartheta_0 = 0$.

The apogee distance and the eccentricity are determined by the formula /29

$$r_\alpha = r_0 \frac{k}{2-k}; \tag{1.36}$$

$$e = k - 1. \tag{1.37}$$

The period of revolution of the satellite

$$T = \frac{2\pi}{\sqrt{\mu}} a^{3/2} = \frac{2\pi}{\sqrt{\mu}} \left(\frac{r_0}{2-k}\right)^{3/2} \tag{1.38}$$

depends only on the orbital semi-axis a.

Selection of Position of Orbital Plane

The position of the orbital plane relative to the Earth is defined by the latitude and longitude of the point of insertion into orbit and by the azimuth of the velocity vector at this point.

The position of the orbital plane relative to the ecliptic and, therefore, relative to stars and the Sun is characterized by the inclination of the orbit to the equatorial plane and the longitude of node Ω_0. The inclination of the orbit is uniquely defined by the azimuth (δ_0) of the velocity vector at the point of insertion and by the latitude (ψ_0) of this point. The longitude of the node is defined by these same quantities, and, in addition, by the time of insertion into orbit. The time of insertion must be taken into account, since the launch is made with Earth rotating in absolute space.

For convenience, let us use sidereal time s which gives directly the angle of rotation of the globe relative to fixed coordinate axes.

Thus, the position of the orbit at any instant of time relative to Earth and Sun is defined by the quantities ψ_0, λ_0, δ_0 and s_0. These quantities are selected based on the purpose of the satellite and its design features, and also several other factors. For example, if the satellite has solar batteries, then it is useful to specifically orient its orbital plane relative to the Sun. If the satellite has a celestial orientation system, these quantities should be selected with reference to the capabilities of its normal functioning.

Let us dwell in more detail on selection of the orbit relative to Earth, which is significant for observation of the satellite, setting up communications with it, and for several other purposes.

Depending on the parameters of the insertion into orbit, the rotation of the Earth has different effects on the initial absolute satellite velocity. Since at any point during the powered section the influence of the Earth's surface velocity can be assumed constant, the increment in velocity in converting from the /30
relative to the absolute coordinate system is

$$\Delta v \approx 0.465 \cos\psi_0 \sin\delta_0 \cos\theta_0 = 0.465 \cos i \cos\theta_0 \quad \text{km/sec},$$

where the numerical coefficient is the Earth's surface velocity at the equator.

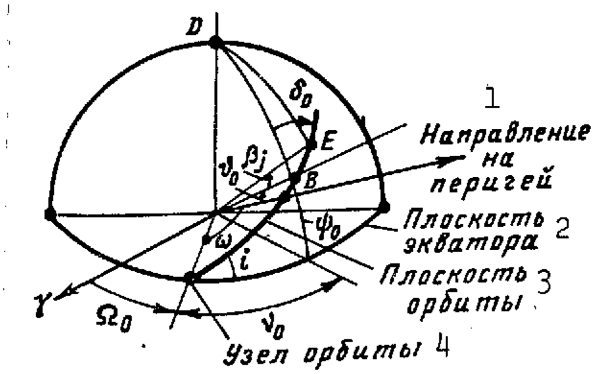


Fig. 1.5. Selecting the position of the orbital plane.

Key: 1. Direction toward perigee; 2. Equatorial plane; 3. Orbital plane; 4. Orbital node

The formula shows that in order to obtain the most distant orbit it is desirable that the latitude of the launch point ψ_0 be selected as close as possible to the equator, and the azimuth δ_0 -- as close as possible to 90°.

Denoting by ψ_j and λ_j the latitude and longitude of point E on Earth, respectively, above which at a given instant t_j the satellite will pass (the satellite is at the zenith over this point at instant t_j), by inspection of the triangle DBE (Fig. 1.5) we can obtain the following formulas:

$$\sin\psi_j = \sin\psi_0\cos\beta_j + \cos\psi_0\sin\beta_j\cos\delta_0; \tag{1.39}$$

$$\lambda_j = \lambda_0 + \Delta\lambda_j - \Delta\lambda_{j\,\mathrm{rot}} \tag{1.40}$$

$$\cos\Delta\lambda_j = \frac{\cos\beta_j - \sin\psi_j\sin\psi_0}{\cos\psi_j\cos\psi_0}; \tag{1.41}$$

$$\sin\Delta\lambda_j = \frac{\sin\beta_j\sin\delta_0}{\cos\psi_j}, \tag{1.42}$$

where $\Delta\lambda_{j\,\mathrm{rot}} = \omega_E(t_j - t_0)$ and ω_E is the absolute angular velocity of rotation of the Earth [rot = rotation; E = Earth].

If the satellite orbit is near-circular, it can be assumed, approximately, that

$$t_j - t_0 = T\frac{\beta_j}{2\pi}. \tag{1.43}$$

When using Eqs. (1.39)-(1.43), we must consider that if the azimuth of the velocity vector relative to a rotating Earth (δ_0') /31
is the initial azimuth, then we must convert to the absolute azimuth δ_0 by the formula

$$\delta_0 \approx \delta_0' + \frac{0.465\cos\psi_0\cos\delta_0}{v\cos\theta_0}. \tag{1.44}$$

Using these formulas, the projections of the orbit onto the Earth were calculated for different azimuths at the point of insertion, whose first orbits, according to the change in angle β_j from 0 to 360°, are shown in Fig. 1.6. In order to obtain the projections of the other orbits, we must shift the curves along the latitude westerly, by the quantity (for the lower orbits)

$$\Delta\lambda_n \approx 15.22(n-1)T^h,$$

where n is the ordinal number of the orbit, and T^h is the period in hours (h).

1.2. General Problem of Inserting Spacecraft into Orbit

Let us examine the case of the insertion into orbit of a satellite by means of some two-stage launch vehicle (LV), whose first stage uses a ramjet engine (RJE) and whose second stage uses a conventional liquid-propellant rocket engine (LPRE). The theory presented below can be easily extended to the case of a large number of launch vehicle stages.

We know that the use of an RJE in the atmospheric section of a flight makes it possible to substantially improve the power capabilities of the LV through use of air oxygen in the Earth's atmosphere as oxidizer.

We will not dwell on methods of calculating RJE, advantages and disadvantages of RJE compared to other types of engines, and the difficulties confronting the builders of this kind of engine. We refer the reader interested in these matters to certain published studies (for example, to [45, 57]).

Our problem will be determining the optimal law of control of the LV engine with RJE over the powered section, that is, determining the optimal control program with respect to the angle of attack α and the mass flow rate of fuel β.

Due to the complexity of the problem we are to solve, it is not possible to obtain analytic functions. Therefore, the method of optimization presented here has been developed as applied to electronic computers.

Formulation of the Problem /33

We will consider the motion of an LV under the following assumptions and presuppositions.

1. The launch vehicle is regarded as a material point.

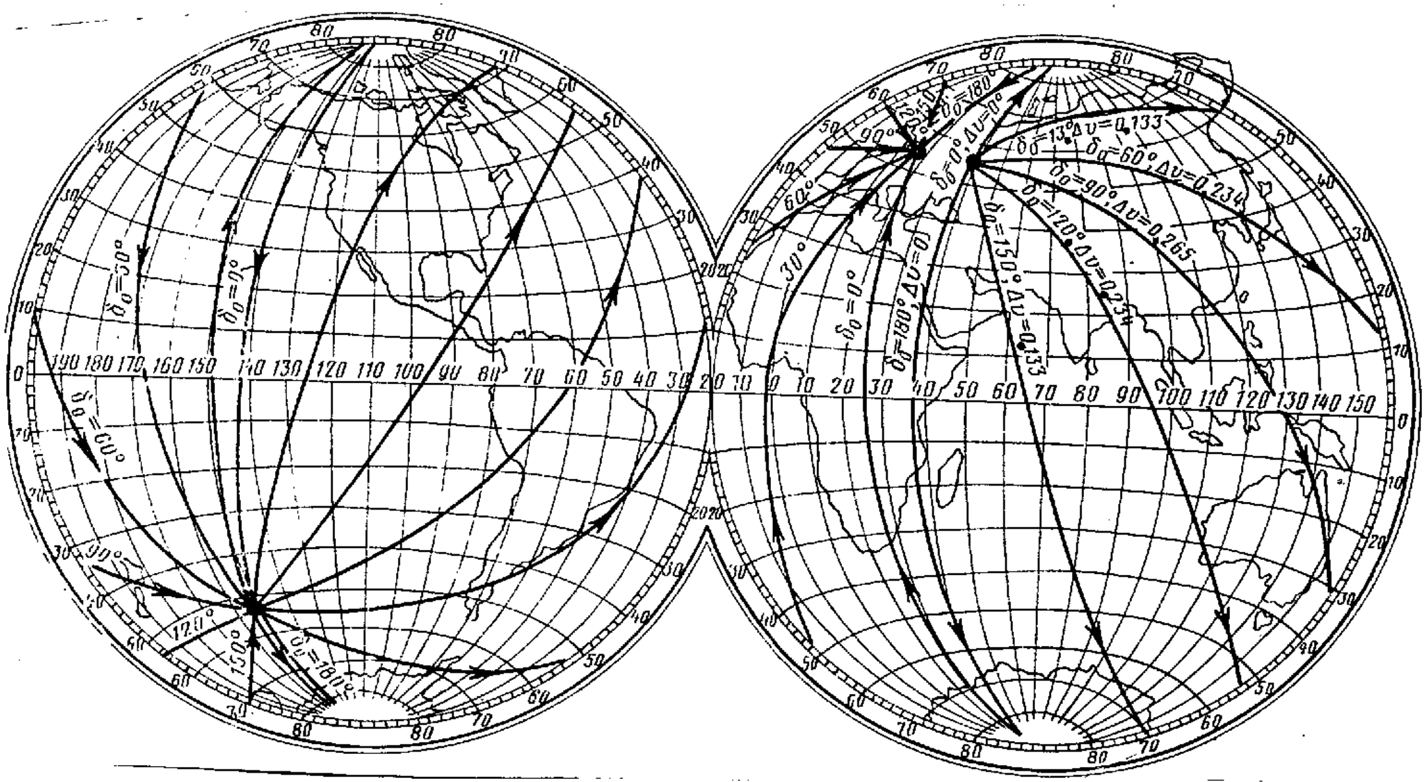


Fig. 1.6. Projections of first orbits for different azimuths δ_0 at point of insertion.

/32

2. The motion of the LV is considered in the vertical plane coinciding with the plane of a great circle of the Earth.

3. The amount of fuel in the first stage, and the total weight of the first and second stages, are given.

4. The field of gravity is central, and the Earth is not rotating.

5. The limiting values of the variation in the angle of attack α_{min} and α_{max} are small, as the result of which we will assume that $\cos \alpha \approx 1$, $\sin \alpha \approx \alpha$ (in the extra-atmospheric section the selection of the boundaries α_{min} and α_{max} will be made by using the terms $P \cos \alpha$ and $P \sin \alpha$ instead of P and $P\alpha$ in the equations of motion (1.45)).

With reference to the foregoing assumptions and presuppositions, the system of equations of motion will become

$$\left.\begin{aligned} \dot{V} &= \frac{P - c_x Sq}{m} - \frac{\mu}{r^2}\sin\theta = f_1; \\ \dot{\theta} &= \frac{P\alpha + c_y Sq}{mV} - \frac{\mu}{r^2 V}\cos\theta + \frac{V\cos\theta}{r} = f_2; \\ \dot{L} &= \frac{r_0}{r} V\cos\theta = f_3; \\ \dot{h} &= V\sin\theta = f_4; \\ \dot{m} &= -\beta = f_5, \end{aligned}\right\} \qquad (1.45)$$

where

$$c_x = c_{x_0}(M) + c_y^{\alpha}(M)\,\alpha^2;\ c_y = c_y^{\alpha}(M)\,\alpha.$$

Let the following constraints be assigned:

$$\xi_1 = (\alpha - \alpha_{min})(\alpha - \alpha_{max}) \leqslant 0; \qquad (1.46)$$

$$\xi_2 = (\beta - \beta_{min})(\beta - \beta_{max}) \leqslant 0; \qquad (1.47)$$

$$\xi_3 = (n_{y_1} - n_{y_1 min})(n_{y_1} - n_{y_1 max}) \leqslant 0; \qquad (1.48)$$

$$\xi_4 = q - q_{max} \leqslant 0; \qquad (1.49)$$

$$\xi_5 = V - l(h) \leqslant 0. \qquad (1.50)$$

Here and above α_{min}, α_{max}, β_{min}, β_{max}, $n_{y\,1\,min}$, $n_{y\,1\,max}$, and q_{max} are some preassigned constant numbers, and $\ell(h)$ is some function of altitude, assigned in advance on the condition that for motion along the curve $V = \ell(h)$, the equilibrium temperature of the LV's surface at some characteristic points on the surface is equal to the allowable maximum value; /34

V is flight velocity
θ is the angle of inclination of the velocity vector to the local horizon
L is flight distance
h is flight altitude
m is instantaneous mass of the LV
P is engine thrust
α is angle of attack
β is mass flow rate of reaction mass
S is characteristic area
q is velocity head
μ $= 3.98 \cdot 10^5$ km^3/sec^2
r is radius-vector
r_0 = 6381 km -- the mean radius of the Earth
c_x is the drag coefficient
c_y is the lift coefficient
n_{y_1} is the transverse load
M is Mach number
P_{sp} is specific thrust.

The inequality (1.46) is a constraint on the angle of attack α, and the inequality (1.47) is a constraint on the mass flow rate of fuel. Since we will take the angle of attack and the mass flow rate of fuel as control function, inequalities (1.46) and (1.47) can be called constraints on the control functions. Inequality (1.48) is a constraint on the transverse g-load n_{y_1}, which as we know is defined by the formula

$$n_{y_1} = \frac{c_y^{\alpha} \alpha S q r^2}{m \mu}.$$

As we can easily see, n_{y_1} depends explicitly on the angle of attack α. However n_{y_1} depends also on the phase coordinates (for example r, m, and so on). Usually this kind of constraint is called constraints on the zero-order phase coordinates.

Inequality (1.49) indicates that the velocity head q is constrained. We can show that $\frac{\partial q}{\partial \alpha} = 0$, $\frac{\partial q}{\partial \beta} = 0$, while at the same time $\frac{\partial}{\partial \alpha}\left(\frac{dq}{dt}\right) \neq 0$. Customarily, this kind of constraint is called a constraint on first-order phase coordinates. The inequality (1.50) is a constraint on the equilibrium temperature of the LV's surface at some characteristic points thereon. It can shown that $\frac{\partial \xi_5}{\partial \alpha} = 0$, $\frac{\partial \xi_5}{\partial \beta} = 0$, a $\frac{\partial}{\partial \alpha}\left(\frac{d\xi_5}{dt}\right) \neq 0$. This indicates that constraint (1.50) is also a constraint on first-order phase coordinates.

Let us take as the optimality criterion the minimum mass flow rate of fuel during the insertion section /35

$$J=\int_{t_0}^{t_K}\beta dt. \tag{1.51}$$

Suppose it is required for the LV to pass from one stage at instant t_0

$$V(t_0)=V_0,\ \theta(t_0)=\theta_0,\ L(t_0)=L_0,\ h(t_0)=h_0,\ m(t_0)=m_0 \tag{1.52}$$

to the final state at instant t_f:

$$V(T_f) = V_f,\ \theta(t_f) = \theta_f,\ L(t_f) = L_f,\ h(t_f) = h_f. \tag{1.53}$$

The problem is formulated as follows: among all the admissible controls α and β satisfying constraints (1.46) and (1.47) and converting the phase point from position (1.52) to position (1.53), we are to find those at which the trajectory V, θ, L, h, and m, corresponding to them and satisfying the differential relations (1.45) would give a minimum for functional (1.51), in order that at any point of the trajectory the inequalities (1.48), (1.49), and (1.50) are satisfied.

We will assume that the functions V, θ, L, and h are continuous and are piecewise-differentiable; the function m is continuous and piecewise-differentiable, and that the moment of staging changes discontinuously by a specific quantity; the functions α and β belong to the class of piecewise-continuous functions; β_{min}, β_{max}, α_{min} and α_{max} are such that they are different for different stages.

Since the LV we are considering is assumed to be of the two-stage type, the thrust, aerodynamic, and other characteristics of each stage differ. This means that this problem is in the category of discontinuous variational problems of the second type.

One feature of this kind of problem is that at the instant of staging the right-hand members of the system (1.45) suffer a discontinuity. We will assume that the surface of the discontinuity is specified by the equation

$$f=m-m_2-\Delta m_1=0, \tag{1.54}$$

where m is the instantaneous mass of the LV; m_2 is the mass of the LV after staging of the first stage; and Δm_1 is the mass of the LV expended in the separation of the first stage.

Method of Optimization

To solve the problem formulated above, let us apply the L.S. Pontryagin principle of the maximum [53]. Under the principle of the maximum, we introduce some function H:

$$H=\sum_{i=1}^{5}\lambda_i f_i, \tag{1.55}$$

where λ_i are certain undetermined multipliers.

Then the first condition for the stationary character of this functional in vector form will become: /36

$$\dot{\lambda}_i=-\frac{\partial H}{\partial x_i}+\psi_3\frac{\partial\xi_3}{\partial x_i}+\psi_4\frac{\partial\dot{\xi}_4}{\partial x_i}+\psi_5\frac{\partial\dot{\xi}_5}{\partial x_i}$$
$$(x_1=V,\ x_2=\theta;\ x_3=L;\ x_4=h;\ x_5=m). \tag{1.56}$$

The multipliers ψ_3, ψ_4, and ψ_5 are defined from the equations

$$\frac{\partial H}{\partial u}=\psi_1\frac{\partial\xi_1}{\partial u}+\psi_2\frac{\partial\xi_2}{\partial u}+\psi_3\frac{\partial\xi_3}{\partial u}+$$
$$+\psi_4\frac{\partial\dot{\xi}_4}{\partial u}+\psi_5\frac{\partial\dot{\xi}_5}{\partial u}$$
$$(u=\alpha,\ \beta), \tag{1.57}$$

where $\dot{\xi}_4$ and $\dot{\xi}_5$ are the total derivatives in time of the functions ξ_4 and ξ_5. Eqs. (1.56) and (1.57) indicate the fact that the desired extremal in the general case is a piecewise-smooth curve consisting of a finite number of sections, each of which either lies at the boundary of a given region (1.46) -- (1.50), or else is within these regions. In the second case, the optimal control is defined from the condition

$$\frac{\partial H}{\partial u}=0. \tag{1.58}$$

In this case all the multipliers $\psi_j = 0$ ($j = 1, 2, 3, 4, 5$).

If the control obtained from Eqs. (1.58) does not satisfy certain constraints (1.46)-(1.50), it is determined from the condition of belonging to these boundaries and the condition of a maximum of the function H. In this case, $\partial H/\partial u \neq 0$. This indicates that one or two multipliers ψ_j are nonzero (the maximum of multipliers ψ_j simultaneously nonzero cannot exceed two, since there are two control functions), and a wholly determined multiplier $\psi_j \neq 0$ occurs only when $\psi_j = 0$ corresponding to it. For example, the optimal control is a control corresponding to the boundary $n_{y_1} = n_{y_1 \max}$, only the multiplier ψ_3 whose numerical value is determined from the solution of Eq. (1.57) is nonzero.

According to Pontryagin's principle for autonomous systems in free time t_f, the exact upper bound to function H can take on the same value, equal to zero, along the entire trajectory, that is, at any instant of time the following equality must be satisfied:

$$H^- = H^+ \tag{1.59}$$

The sign "-" indicates that the function H is taken at the left, and "+" is taken at the right of the fixed point at the given instant of time. /37

As for the multipliers λ_i, at the moment the first-order phase coordinates arrive at the bounds of the admissible regions, and also at the instant of staging, they can suffer a discontinuity of the first kind.

The conditions for a discontinuity of the multipliers λ_i at the instant of insertion are as follows:

-- at the boundary $\psi_4 = 0$

$$\lambda_i^+ = \lambda_i^- + \nu \frac{\partial \xi_4}{\partial x_i}; \tag{1.60}$$

-- at the boundary $\psi_5 = 0$

$$\lambda_i^+ = \lambda_i^- + \nu_1 \frac{\partial \xi_5}{\partial x_i}. \tag{1.61}$$

At the moment of staging, the conditions of discontinuity of the cofactors are of the form

$$\lambda_i^+ = \lambda_i^- + \nu_2 \frac{\partial f}{\partial x_i}. \tag{1.62}$$

Conjugate System of Differential Equations

As already indicated above, the characteristics of each stage differ. Because of this, the trajectory breaks down into individual stages separated by the surface (1.50), in each of which stages the function H is of the form

$$H = \lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3 + \lambda_4 f_4 + \lambda_5 f_5. \tag{1.63}$$

Eqs. (1.56) in phase coordinates at each of these stages will be:

$$\left.\begin{aligned}
\dot{\lambda}_1 &= -\lambda_1\frac{\partial f_1}{\partial V}-\lambda_2\frac{\partial f_2}{\partial V}-\lambda_3\frac{\partial f_3}{\partial V}-\lambda_4\frac{\partial f_4}{\partial V}+\psi_3\frac{\partial \dot{\xi}_3}{\partial V}+\psi_4\frac{\partial \dot{\xi}_4}{\partial V}+\psi_5\frac{\partial \dot{\xi}_5}{\partial V};\\
\dot{\lambda}_2 &= -\lambda_1\frac{\partial f_1}{\partial \theta}-\lambda_2\frac{\partial f_2}{\partial \theta}-\lambda_3\frac{\partial f_3}{\partial \theta}-\lambda_4\frac{\partial f_4}{\partial \theta}+\psi_4\frac{\partial \dot{\xi}_4}{\partial \theta}+\psi_5\frac{\partial \dot{\xi}_5}{\partial \theta};\\
\dot{\lambda}_3 &= 0;\\
\dot{\lambda}_4 &= -\lambda_1\frac{\partial f_1}{\partial h}-\lambda_2\frac{\partial f_2}{\partial h}-\lambda_3\frac{\partial f_3}{\partial h}+\psi_3\frac{\partial \dot{\xi}_3}{\partial h}+\psi_4\frac{\partial \dot{\xi}_4}{\partial h}+\psi_5\frac{\partial \dot{\xi}_5}{\partial h};\\
\dot{\lambda}_5 &= -\lambda_1\frac{\partial f_1}{\partial m}-\lambda_2\frac{\partial f_2}{\partial m}+\psi_3\frac{\partial \dot{\xi}_3}{\partial m}+\psi_4\frac{\partial \dot{\xi}_4}{\partial m}+\psi_5\frac{\partial \dot{\xi}_5}{\partial m}.
\end{aligned}\right\} \qquad (1.64)$$

Here

$$\frac{\partial f_1}{\partial V}=\frac{1}{m}\left(\frac{\partial P}{\partial V}-\frac{\partial c_x}{\partial V}Sq-c_xS\varrho V\right);$$

$$\frac{\partial f_1}{\partial \theta}=-\frac{\mu}{r^2}\cos\theta;$$

$$\frac{\partial f_1}{\partial h}=\frac{1}{m}\left(\frac{\partial P}{\partial h}-\frac{1}{2}c_xSV^2\frac{\partial \varrho}{\partial h}\right)+\frac{2\mu}{r^3}\sin\theta;$$ /38

$$\frac{\partial f_1}{\partial m}=-\frac{1}{m^2}(P-c_xSq);$$

$$\frac{\partial f_2}{\partial V}=\frac{1}{mV}\left(\frac{\partial P}{\partial V}\alpha+\frac{\partial c_y}{\partial V}Sq+c_yS\varrho V\right)-\frac{1}{mV^2}(P\alpha+c_ySq)+$$
$$+\frac{\mu}{r^2V^2}\cos\theta+\frac{\cos\theta}{r};$$

$$\frac{\partial f_2}{\partial \theta}=\frac{\mu}{r^2V}\sin\theta-\frac{V\sin\theta}{r};$$

$$\frac{\partial f_2}{\partial h}=\frac{1}{mV}\left(\frac{\partial P}{\partial h}\alpha+\frac{1}{2}c_ySV^2\frac{\partial \varrho}{\partial h}\right)+\frac{2\mu}{r^3V}\cos\theta-\frac{V\cos\theta}{r^2};$$

$$\frac{\partial f_2}{\partial m}=-\frac{P\alpha+c_ySq}{m^2V};$$

$$\frac{\partial f_3}{\partial V}=\frac{r_0}{r}\cos\theta;$$

$$\frac{\partial f_3}{\partial \theta}=-\frac{r_0}{r}V\sin\theta;$$

$$\frac{\partial f_3}{\partial h}=-\frac{r_0}{r^2}V\cos\theta;$$

$$\frac{\partial f_4}{\partial V}=\sin\theta;$$

$$\frac{\partial f_4}{\partial \theta}=V\cos\theta;$$

$$\frac{\partial \dot{\xi}_3}{\partial V}=\frac{Sr^2}{m\mu}(2n_{y_1}-n_{y_1\min}-n_{y_1\max})\left(\frac{\partial c_y}{\partial V}q+c_y\varrho V\right);$$

$$\frac{\partial \dot{\xi}_3}{\partial h}=\frac{Sr^2}{m\mu}(2n_{y_1}-n_{y_1\min}-n_{y_1\max})\left(\frac{1}{2}c_yV^2\frac{\partial\varrho}{\partial h}+\frac{2}{r}c_yq\right);$$

$$\frac{\partial \dot{\xi}_3}{\partial m}=-\frac{c_ySqr^2}{m^2\mu}(2n_{y_1}-n_{y_1\min}-n_{y_1\max});$$

$$\dot{\xi}_4=\frac{1}{2}\frac{\partial\varrho}{\partial h}V^3\sin\theta+\varrho Vf_1;$$

$$\frac{\partial \dot{\xi}_4}{\partial V}=\frac{3}{2}\frac{\partial\varrho}{\partial h}V^2\sin\theta+\varrho f_1+\varrho V\frac{\partial f_1}{\partial V};$$

$$\frac{\partial \dot{\xi}_4}{\partial \theta}=\frac{1}{2}\frac{\partial\varrho}{\partial h}V^3\cos\theta+\varrho V\frac{\partial f_1}{\partial \theta};$$

$$\frac{\partial \dot{\xi}_4}{\partial h}=\frac{1}{2}\frac{\partial^2\varrho}{\partial h^2}V^3\sin\theta+\frac{\partial\varrho}{\partial h}Vf_1+\varrho V\frac{\partial f_1}{\partial h};$$ /39

$$\frac{\partial \dot{\xi}_4}{\partial m}=\varrho V\frac{\partial f_1}{\partial m};$$

$$\dot{\xi}_5=f_1-\frac{\partial l(h)}{\partial h}V\sin\theta;$$

$$\frac{\partial \dot{\xi}_5}{\partial V}=\frac{\partial f_1}{\partial V}-\frac{\partial l(h)}{\partial h}\sin\theta;$$

$$\frac{\partial \dot{\xi}_5}{\partial \theta}=\frac{\partial f_1}{\partial \theta}-\frac{\partial l(h)}{\partial h}V\cos\theta;$$

$$\frac{\partial \dot{\xi}_5}{\partial h}=\frac{\partial f_1}{\partial h}-\frac{\partial^2 l(h)}{\partial h^2}V\sin\theta;$$

$$\frac{\partial \dot{\xi}_5}{\partial m}=\frac{\partial f_1}{\partial m};$$

$$\frac{\partial c_x}{\partial V}=\frac{1}{a^*}\left(\frac{\partial c_{x_0}}{\partial \mathrm{M}}+\frac{\partial c_y^\alpha}{\partial \mathrm{M}}\alpha^2\right);$$

$$\frac{\partial c_y}{\partial V}=\frac{1}{a^*}\frac{\partial c_y^\alpha}{\partial \mathrm{M}}\alpha.$$

As indicated earlier, the LV we are considering is a two-stage LV using a RJE in the first stage, and a LPRE in the second.

As we know, the specific thrust of an RJE for a selected engine and specified fuel components depends basically on the actual excess air coefficient α_Φ and the Mach number M, that is, it can be approximated by a polynomial of the n-th degree

$$P_{sp}=A_0(M)+A_1(M)\alpha_\Phi+A_2(M)\alpha_\Phi^2+\ldots+A_n(M)\alpha_\Phi^n, \tag{1.65}$$

where

$$a_{\phi}=\frac{D_0 k_1}{\beta};\quad D_0=\frac{\varrho V F\varphi}{\varkappa};\ k_1=a(M)\alpha+b(M);$$

$A_0(M)$, $A_1(M)$, $A_2(M)$, ..., $A_n(M)$, $a(M)$, $b(M)$ are some functions of the M number; k_1 is the coefficient of flow compression in the LV body; F is the area of the diffuser inlet; ϕ is the coefficient of diffuser mass flow; ρ is the atmospheric density; and κ is the stoichiometric ratio.

Let us limit ourselves to examining only the first two terms of series (1.65). Then the thrust P will be of the form

$$P=g_0[A_1D_0(a\alpha+b)+A_0\beta]. \tag{1.66}$$

The thrust of a LPRE, as we know, is determined by the formula /40

$$P=g_0(J_0\beta-e_0S_a\varrho), \tag{1.67}$$

where g_0 = 9.81 m/sec^2 is the equation of terrestrial gravity at sea level; J_0 is the specific thrust of an LPRE (in vacuo); S_α is the area of the nozzle discharge; and $e_0 = RT/g_0$ (R = 287.05 m^2/ /sec^2·deg is the gas constant per kg-mass of air, and T is air temperature).

Then the partial derivatives in P will be:

-- over the section up to the separation of the first stage

$$\frac{\partial P}{\partial V}=g_0\left\{A_1D_0\left[\left(\frac{1}{V}+\frac{1}{a^*\varphi}\frac{\partial\varphi}{\partial M}+\frac{1}{a^*A_1}\frac{\partial A_1}{\partial M}\right)(a\alpha+b)+\right.\right.$$
$$\left.\left.+\frac{1}{a^*}\left(\frac{\partial a}{\partial M}\alpha+\frac{\partial b}{\partial M}\right)\right]+\frac{\beta}{a^*}\frac{\partial A_0}{\partial M}\right\};$$
$$\frac{\partial P}{\partial h}=g_0A_1D_0\frac{1}{\varrho}\frac{\partial\varrho}{\partial h}(a\alpha+b);$$

-- over the section after separation of the first stage:

$$\frac{\partial P}{\partial V}=0;$$
$$\frac{\partial P}{\partial h}=-g_0e_0S_a\frac{\partial\varrho}{\partial h}.$$

In solving Eqs. (1.64), we everywhere assumed that the speed of sound a* is constant in altitude. It was also assumed that the functions ρ, $l(h)$, ϕ, c^α, c_{x0}, a, b, A_0, and A_1, and also their partial derivatives $\frac{\partial\varrho}{\partial h}, \frac{\partial^2\varrho}{\partial h^2}, \frac{\partial l(h)}{\partial h}, \frac{\partial^2 l(h)}{\partial h^2}; \frac{\partial\gamma}{\partial M}, \frac{\partial c_{x0}}{\partial M}, \frac{\partial c_y^\alpha}{\partial M}, \frac{\partial a}{\partial M}, \frac{\partial b}{\partial M}, \frac{\partial A_0}{\partial M}$, and $\frac{\partial A_1}{\partial M}$ are known in advance and specified in the form of tables or appropriate graphs.

Determination of Optimal Control

At any fixed moment of time τ, the phase coordinates x_i and multipliers (cofactors) λ_i are fixed. Therefore the function (1.63) at this instant depends only on the controls α and β. We also need to find α and β such that the function H takes on its exact upper bound.

We will show that the constraint on the transverse g-load n_{y_1} can be reduced to a constraint on the angle of attack α.

Actually, in order for the transverse g-load n_{y_1} at any instant of time τ not to exceed its specified value, it is necessary that the angle of attack α lie in the interval /41

$$\alpha_{1\min} \leqslant \alpha \leqslant \alpha_{1\max},$$

where

$$\alpha_{1\min} = \frac{n_{y_1\min} m\mu}{c_y^\alpha S q r^2}; \quad \alpha_{1\max} = \frac{n_{y_1\max} m\mu}{c_y^\alpha S q r^2}.$$

Therefore, it can be assumed that at any fixed moment two constraints are imposed on the control function α:

$$\alpha_{\min} \leqslant \alpha \leqslant \alpha_{\max}; \quad \alpha_{1\min} \leqslant \alpha \leqslant \alpha_{1\max}.$$

The last two inequalities can be replaced by the single inequality of the form

$$\alpha_{2\min} \leqslant \alpha \leqslant \alpha_{2\max}, \tag{1.68}$$

where

$$\left.\begin{aligned} &\alpha_{2\min} = \alpha_{\min}, \text{ if } \alpha_{\min} \geqslant \alpha_{1\min};\\ &\alpha_{2\min} = \alpha_{1\min}, \text{ if } \alpha_{\min} < \alpha_{1\min};\\ &\alpha_{2\max} = \alpha_{\max}, \text{ if } \alpha_{\max} \leqslant \alpha_{1\max};\\ &\alpha_{2\max} = \alpha_{1\max}, \text{ if } \alpha_{\max} > \alpha_{1\max}. \end{aligned}\right\} \tag{1.69}$$

The function H can be represented in the form:

-- over the section up to the separation of the first stage

$$H = A_{11}\alpha^2 + E_{11}\alpha + F_{11}\alpha\beta + N_{11}\beta + T_{11}; \tag{1.70}$$

-- over the section after the separation of the first stage

$$H = H_{12}\alpha^2 + E_{12}\alpha + F_{12}\alpha\beta + N_{12}\beta + T_{12}. \tag{1.71}$$

Here

$$A_{11} = \frac{1}{m}\left(\frac{\lambda_2}{V} g_0 A_1 D_0 a - \lambda_1 c_y^\alpha S q\right);$$

$$E_{11} = \frac{\lambda_1}{m} g_0 A_1 D_0 a + \frac{\lambda_2}{mV}\left(g_0 A_1 D_0 b + c_y^\alpha S q\right);$$

$$F_{11} = \frac{\lambda_2}{mV} g_0 A_0;$$

$$N_{11} = \frac{\lambda_1}{m} g_0 A_0 - \lambda_5;$$

$$T_{11} = \lambda_1\left(\frac{g_0 A_1 D_0 b - c_{x0} S q}{m} - \frac{\mu}{r^2}\sin\theta\right) + \frac{\lambda_2}{r}\cos\theta\left(V - \frac{\mu}{rV}\right) + \lambda_3 \frac{r_0}{r} V \cos\theta + \lambda_4 V \sin\theta;$$

$$A_{12} = -\frac{\lambda_1}{m} c_y^\alpha S q;$$

/42

$$E_{12} = \frac{\lambda_2}{mV}\left(c_y^\alpha S q - g_0 e_0 S_a \varrho\right);$$

$$F_{12} = \frac{\lambda_2}{mV} g_0 J_0;$$

$$N_{12} = \frac{\lambda_1}{m} g_0 J_0 - \lambda_5;$$

$$T_{12} = -\lambda_1\left(\frac{g_0 e_0 S_a \varrho + c_{x0} S q}{m} + \frac{\mu}{r^2}\sin\theta\right) + \frac{\lambda_2}{r}\cos\theta\left(V - \frac{\mu}{rV}\right) + \lambda_4 V \sin\theta + \lambda_3 \frac{r_0}{r} V \cos\theta.$$

Similarly, let us represent the relations imposed on the trajectory for motion along the boundaries $\xi_4 = 0$ and $\xi_5 = 0$. For the LV to move along these boundaries, the following must obtain: $\dot{\xi}_4 = 0$ and $\dot{\xi}_5 = 0$.

In expanded form, this can be written as:

-- over the section prior to separation of the first stage

$$\dot{\xi}_4 = A_{21}\alpha^2 + E_{21}\alpha + N_{21}\beta + T_{21} = 0; \quad (1.72)$$

$$\dot{\xi}_5 = A_{31}\alpha^2 + E_{31}\alpha + N_{31}\beta + T_{31} = 0. \quad (1.73)$$

-- over the section after separation of the first stage

$$\dot{\xi}_4 = A_{22}\alpha^2 + E_{22}\alpha + N_{22}\beta + T_{22} = 0; \quad (1.74)$$

$$\dot{\xi}_5 = A_{32}\alpha^2 + E_{32}\alpha + N_{32}\beta + T_{32} = 0. \quad (1.75)$$

Here

$$A_{21} = -\frac{\varrho V}{m} c_y^{\alpha} Sq; \quad E_{21} = \frac{\varrho V}{m} g_0 A_1 D_0 a;$$

$$N_{21} = \frac{\varrho V}{m} g_0 A_0;$$

$$T_{21} = \frac{1}{2} \frac{\partial \varrho}{\partial h} V^3 \sin\theta + \varrho V \left(\frac{g_0 A_1 D_0 b - c_{x0} Sq}{m} - \frac{\mu}{r^2} \sin\theta \right);$$

$$A_{31} = -\frac{c_y^{\alpha} Sq}{m}; \quad E_{31} = \frac{g_0 A_1 D_0 a}{m}; \quad N_{31} = \frac{g_0 A}{m};$$

$$T_{31} = \frac{g_0 A_1 D_0 b - c_{x0} Sq}{m} - \frac{\mu}{r^2} \sin\theta - \frac{\partial l(h)}{\partial h} V \sin\theta;$$

$$A_{22} = A_{21}; \quad E_{22} = 0; \quad N_{22} = \frac{\varrho V}{m} g_0 J_0;$$

$$T_{22} = \frac{1}{2} \frac{\partial \varrho}{\partial h} V^3 \sin\theta - \varrho V \left(\frac{g_0 e_0 S_a \varrho + c_{x0} Sq}{m} + \frac{\mu}{r^2} \sin\theta \right);$$

$$A_{32} = A_{31}; \quad E_{32} = 0; \quad N_{32} = \frac{g_0 J_0}{m};$$

$$T_{32} = -\frac{g_0 e_0 S_a \varrho + c_{x0} Sq}{m} - \frac{\mu}{r^2} \sin\theta - \frac{\partial l(h)}{\partial h} V \sin\theta.$$

Let us present an example of the selection of the optimal control α and β at some fixed instant of time τ, where we will assume that at this instant no staging occurs. The selection of the optimal control at the instant of staging will be considered a little later. /43

As we can see Eqs. (1.70) and (1.71) are identical with respect to variables α and β and differ only in the coefficients appearing by these variables. These coefficients at any instant of time (except for the instant of staging) are uniquely defined. As a result, to condense the notation, we replace Eqs. (1.70) and (1.71) with a single formula of the form

$$H = A\alpha^2 + E\alpha + F\alpha\beta + N\beta + T. \quad (1.76)$$

The coefficients A, E, F, and F of these formulas are equal to the following:

-- over the section before separation of the first stage

$$A=A_{11};\ E=E_{11};\ F=F_{11};\ N=N_{11};\ T=T_{11};$$

-- over the section after separation of the first stage

$$A=A_{12};\ E=E_{12};\ F=F_{12};\ N=N_{12};\ T=T_{12}.$$

It is required to find in Eq. (1.76) α and β satisfying the inequalities

$$\left.\begin{array}{l}\alpha_{2\min}\leqslant\alpha\leqslant\alpha_{2\max};\\ \beta_{\min}\leqslant\beta\leqslant\beta_{\max},\end{array}\right\} \quad (1.77)$$

where the function H would take on its largest value.

Let us examine the case when the constraints (1.49) and (1.50) have been satisfied. We will show that the function H takes on its largest value only along the boundaries $\beta = \beta_{min}$ and $\beta = \beta_{max}$.

Let us set up the determinant:

$$\Delta=\begin{vmatrix}\frac{\partial^2 H}{\partial\alpha^2} & \frac{\partial^2 H}{\partial\alpha\partial\beta}\\ \frac{\partial^2 H}{\partial\alpha\partial\beta} & \frac{\partial^2 H}{\partial\beta^2}\end{vmatrix}=\begin{vmatrix}2A & F\\ F & 0\end{vmatrix}=-F^2\leqslant 0. \quad (1.78)$$

This shows that the function H can have within the specified region [domain] (1.77) only a saddle-point.

Since the mass flow rate of β appears linearly in the function of H, this means that along the boundaries $\alpha = \alpha_{2\,min}$ and $\alpha = \alpha_{2\,max}$, the function H takes on its largest value either when $\beta = \beta_{min}$ or when $\beta = \beta_{max}$, that is, the maximum of the function H can be realized only at the apices of the rectangle (1.77). Thus, the maximum of the function H can be realized only along the boundaries $\beta = \beta_{min}$ and $\beta = \beta_{max}$.

We will present points at which the extremum of the function (1.76) can be realized as the LV moves within the domains (1.49) and (1.50), in final form: /44

$$\left.\begin{array}{l}1)\ \alpha_{2\min}\quad \beta_{\min};\\ 2)\ \alpha_{2\max}\quad \beta_{\min};\\ 3)\ \alpha_{2\min}\quad \beta_{\max};\\ 4)\ \alpha_{2\max}\quad \beta_{\max};\\ 5)\ \alpha_1=-\frac{E+F\beta_{\min}}{2A}\quad \beta_{\min};\\ 6)\ \alpha_2=-\frac{E+F\beta_{\max}}{2A}\quad \beta_{\max}.\end{array}\right\} \quad (1.79)$$

To condense the notation for defining the optimal control as the LV moves along the boundaries (1.49) and (1.50), let us introduce some polynomial

$$\gamma=B_0\alpha^2+B_1\alpha+B_2\beta+B_3=0. \quad (1.80)$$

The coefficients B_0, B_1, B_2, and B_3 of the polynomial are taken as equal to the following:

-- over the section before separation of the first stage:

a) for motion along the boundary $\xi_4 = 0$

$$B_0=A_{21};\ B_1=E_{21};\ B_2=N_{21};\ B_3=T_{21};$$

b) for motion along the boundary $\xi_5 = 0$

$$B_0=A_{31};\ B_1=E_{31};\ B_2=N_{31};\ B_3=T_{31};$$

-- along the section after separation of the first stage

a) for motion along the boundary $\xi_4 = 0$

$$B_0=A_{22};\ B_1=E_{22};\ B_2=N_{22};\ B_3=T_{22};$$

b) for motion along the boundary $\xi_5 = 0$

$$B_0=A_{32};\ B_1=E_{32};\ B_2=N_{32};\ B_3=T_{32}.$$

Polynomial (1.80) is a generalization of the polynomials (1.72), (1.73), (1.75), and (1.75) we have already introduced.

By determining β from this polynomial and inserting it into Eq. (1.76), we will have

$$H=B_4\alpha^3+B_5\alpha^2+B_6\alpha+B_7, \quad (1.81)$$

where

$$B_4 = -\frac{FB_0}{B_2};\; B_5 = A - \frac{1}{B_2}(FB_1 + NB_0);$$
$$B_6 = E - \frac{1}{B_2}(FB_3 + NB_1);\; B_7 = T - \frac{NB_3}{B_2}.$$

Then we can quite readily find the points at which the function H can have an extremum for motion of the LV along the bounds (1.49) and (1.50). They will be: /45

$$\left.\begin{array}{lll}
1)\ \alpha_3 = \dfrac{-B_5 - \sqrt{B_5^2 - 3B_4B_6}}{3B_4} & \beta_1 = -\dfrac{B_0\alpha_3^2 + B_1\alpha_3 + B_3}{B_2}; \\
2)\ \alpha_4 = -\dfrac{B_1 + \sqrt{B_1^2 - 4B_0(B_2\beta_{min} + B_3)}}{2B_0} & \beta_{min}; \\
3)\ \alpha_5 = \dfrac{-B_1 + \sqrt{B_1^2 - 4B_0(B_2\beta_{min} + B_3)}}{2B_0} & \beta_{min}; \\
4)\ \alpha_6 = -\dfrac{B_1 + \sqrt{B_1^2 - 4B_0(B_2\beta_{max} + B_3)}}{2B_0} & \beta_{max}; \\
5)\ \alpha_7 = \dfrac{-B_1 + \sqrt{B_1^2 - 4B_0(B_2\beta_{max} + B_3)}}{2B_0} & \beta_{max}; \\
6)\ \alpha_{2min} & \beta_2 = -\dfrac{B_0\alpha_{2min}^2 + B_1\alpha_{2min} + B_3}{B_2}; \\
7)\ \alpha_{2max} & \beta_3 = -\dfrac{B_0\alpha_{2max}^2 + B_1\alpha_{2max} + B_3}{B_2}.
\end{array}\right\} \quad (1.82)$$

The values of α_1, α_2, α_3, α_4, α_5, α_6, and α_7, and also β_1, β_2, and β_3 obtained by Eqs. (1.79) and (1.82) must satsify the corresponding conditions (1.77). If for each point these conditions are satisfied for both α and β, and, in addition, the conditions (1.49) and (1.50) are satisfied, the resulting point is left for further investigation; but if these conditions are not satisfied or if one of the values of α or β are imaginary, it is discarded.

It should be noted that at any instant of time the number of points satisfying these conditions is smaller than is presented here. Actually, many of these points preclude each other. Some of them are imaginary. However, we cannot in advance say exactly to which point at a given instant of time the optimal control corresponds. Therefore, we must check all these points and leave those which satisfy our requirements.

In the following we will not qualify this in every instance, and if we consider a particular point, we will assume that the conditions specified for it have been met.

By calculating the values of the function (1.76) for points satisfying all the above-listed requirements, let us select the point at which the function H takes on its exact upper bound. The control α and β thus obtained is the optimal control at any fixed instant of time τ (except for the instant of staging). /46

Determining Optimal Control at Instant of Staging

As already indicated earlier, at the instant of staging β_{min}, β_{max}, S, P_{sp}, c_{x_0}, α_{min}, and α_{max} can suffer discontinuities, as can other LV parameters; the Lagrangian also suffer discontinuities.

We are required, speaking figuratively, to "dovetail" two different sections: the section prior to staging and the section after staging. Condition (1.59) will be the criterion of this "dovetailing," and this condition must be realized for an exact upper bound of function H^+ in terms of control.

From conditions (1.62), we have:

$$\begin{gathered}\lambda_i^+ = \lambda_i^- \ (i = 1,\ 2,\ 3,\ 4),\\ \lambda_5^+ = \lambda_5^- + \nu_2,\end{gathered} \tag{1.83}$$

that is, at the instant of staging only the multiplier λ_5 suffers a discontinuity, and the rest are continuous.

The function H^+ at the instant of staging is of the form (1.71). All the coefficients of this function, with the exception of N_{12}, are known at this instant, since the characteristics of the LV after staging are specified ahead of time, and the multipliers λ_i (i = 1, 2, 3, 4) at the instant of staging are continuous. As for the coefficient N_{12}, because the multiplier λ_5 has a discontinuity, we as yet do not know this coefficient.

Let us note that the formula functions (1.79) and (1.82) serving for the determination of the optimal control within the sections separated by surface (1.54) are valid also at the instant of staging, and in those functions where the coefficient N does not appear, this control is uniquely defined both at the left and at the right, and it requires only that into these functions the LV parameters corresponding to the given stage be substituted.

As we can see from Eq. (1.79), the coefficient N does not appear in the formulas for the optimal control, and in the functions (1.82) it appears only in formula 1).

Let us define the optimal control and thus the coefficient N_{12} for the case when staging occurs within the domains (1.49) and (1.50).

Let us write the condition (1.59) at the instant of staging in expanded form:

$$H^- = A_{11}\alpha^{-2} + E_{11}\alpha^- + F_{11}\alpha^-\beta^- + N_{11}\beta^- + T_{11} =$$
$$= A_{12}\alpha^{+2} + E_{12}\alpha^+ + F_{12}\alpha^+\beta^+ + N_{12}\beta^+ + T_{12} = H^+. \quad (1.84)$$

Since H^- at the instant of staging is known, condition (1.84) can be represented as: /47

$$A_{12}\alpha^{+2} + E_{12}\alpha^+ + T_{12} - H^- + (F_{12}\alpha^+ + N_{12})\beta^+ = 0. \quad (1.85)$$

We note that only one of the limiting values β_{min} or β_{max} can be the optimal value β^+, just as β^- (with the only difference that their numerical values [β_{min} and β_{max}] correspond to the second stage and can differ from the numerical values of β_{min} and β_{max} at the first stage).

Let us denote the part of the function (1.85) independent of β^+ by C, that is

$$C = A_{12}\alpha^{+2} + E_{12}\alpha^+ + T_{12} - H^-. \quad (1.86)$$

Let us substitute here the values α^+ corresponding to points 1, 2 and 5 from formulas (1.79).

We note that the value of β^+ corresponding to point 5 is substituted into formula (1.86) only if it satisfies condition (1.77), and the coefficient $A_{12} < 0$. It is not considered in the remaining cases.

Since the points 1, 2 and 5 correspond to $\beta^+ = \beta_{min}$, this means that the following must obtain at these points:

$$F_{12}\alpha^+ + N_{12} < 0, \quad (1.87)$$

or, because $\beta_{min} > 0$, C must be greater than zero.

Let us select from points 1, 2, and 5 those for which $C > 0$, and let us reject the remaining.

Let us substitute the remaining points into the formula

$$C_1 = C + F_{12}\alpha^+\beta^+ \tag{1.88}$$

and, let us dwell on the points for which C_1 is at a maximum. The coefficient N_{12} corresponding to the points thus selected is defined by the formula

$$N_{12} = \frac{H^{+} - A_{12}\alpha^{+2} - E_{12}\alpha^{+} - F_{12}\alpha^{+}\beta^{+} - T_{12}}{\beta^{+}}. \tag{1.89}$$

Let us substitute into Eq. (1.86) the values α^+ corresponding to points 3, 4, and 6 of formulas (1.79), where the latter point is substituted into this formula only if the conditions (1.77) and $A_{12} < 0$ are satisfied for it. Let us select from these points those for which $C < 0$, and let us reject the rest.

We substitute the remaining points into Eq. (1.88) and we dwell on the point at which C_1 has a maximum. The coefficient N_{12} corresponding to this point is defined by formula (1.89).

By carrying out this study, we find at each of the bounds $\beta = \beta_{min}$ and $\beta = \beta_{max}$ one point each (the case when there is no point of this kind along any of the bounds is possible), at which the exact upper bound of function H^+ can be realized.

Now we must select from these two points the one for which, at the coefficient N_{12} found for it the function H^+ realizes its exact upper bound at this point. /48

Obviously, if we take as the optimal control the point for which the coefficient N_{12} is the smallest, we obtain the result that at any other point for the coefficient N_{12} thus determined the function H^+ is smaller than at the point we selected.

Thus defining the optimal control α^+ and β^+ and the coefficient N_{12}, we can also find the multiplier λ_5^+:

$$\lambda_5^+ = \frac{\lambda_1^-}{m^+} g_0 I_0 - N_{12}. \tag{1.90}$$

Now let us show that the control thus selected always exists.

If $A_{12} < 0$ and α_1^+, α_2^+ (see Eq. (1.79)) satisfies the condition (1.77), by substituting α_1^+ and α_2^+ into Eq. (1.86) we can see that $C(\alpha_1^+) > C(\alpha_2^+)$. Hence it follows that if $C(\alpha_1^+) < 0$ (the case when the optimal control is absent along the bound $\beta^+ = \beta_{min}$, then $C(\alpha_2^+) < 0$ and this means that the optimal control exists along the bound $\beta^+ = \beta_{max}$. If $C(\alpha_2^+) > 0$ (the case when there is not optimal control along the bound $\beta^+ = \beta_{max}$), then $C(\alpha_1^+) > 0$ and this means that the optimal control exists along the bound $\beta^+ = \beta_{min}$.

If $A_{12} > 0$ (or $A_{12} < 0$, but α_1^+ and α_2^+ do not satisfy condition (1.77)), obviously the function H^+ realizes its maximum along the boundaries $\alpha^+ = \alpha_{2\,min}$ and $\alpha^+ = \alpha_{2\,max}$. Substituting into Eq. (1.86) the values of α^+, we uniquely define the sign of C, and hence thus the values of β^+ corresponding to the bounds. Therefore, in this case there is an optimal control.

Now let us assume that $A_{12} < 0$, and that one of the values of α_1^+ or α_2^+ (for example, α_1^+) does not satisfy condition (1.77). Obviously, if $\alpha_1^+ > \alpha_{2\,max}$, then we must select $\alpha_{2\,max}$; but if $\alpha_1^+ < \alpha_{2\,min}$, we must take $\alpha_{2\,min}$. Let us denote this control by the symbol α_9. Since α_2^+ satisfies condition (1.88) and α_1^+ does not, then clearly α_9 is in this segment $[\alpha_1^+, \alpha_2^+]$. On the other hand, the maximum of the function C (see Eq. (1.86)) is realized when $\alpha_8 = -E_{12}/2A_{12}$. By comparing the formulas for α_1^+ and α_2^+ (see Eq. (1.79)) with the formula for α_8, we see that the value of α_8 lies outside the segment $[\alpha_1^+, \alpha_2^+]$. This means that in this segment the function C either decreases or increases with respect to α. We earlier showed that $C(\alpha_1^+) > C(\alpha_2^+)$. Now we can write that $C(\alpha_1^+) > C(\alpha_9) > C(\alpha_2^+)$. Further considerations are analogous to the case when $A_{12} < 0$ and α_1^+, α_2^+ satisfy condition (1.77).

Thus, at the instant of staging the optimal control always exists and does so uniquely (an exception may be found in several /49
singular cases, which we do not consider here).

Now let us define the optimal control and thus the coefficient N_{12} for the case when staging occurs along the bounds (1.49) and(1.50).

As already indicated, in the formulas 1) of the function (1.82) there appears the unknown coefficient N_{12}. So α_3 and thus β_1 cannot be determined directly from these formulas.

We can see that at the instant of staging the coefficient B_4^+ and B_6^+ do not depend on the coefficient N_{12}, while B_5^+ and B_7^+ are functions of N_{12}. Let us represent B_5^+, N_{12}, and B_7^+ as functions of the angle of attack α_3^+, and the coefficients B_4^+ and B_6^+ (this can be done by using formula 1) of functions (1.82) and the coefficients earlier obtained for the formula (1.81)).

$$B_5^+ = -\frac{3B_4^+\alpha_3^{+2} + B_6^+}{2\alpha_3^+}; \tag{1.91}$$

$$N_{12} = \frac{3\left(B_4^+\alpha_3^{+2} + 2A_{12}\alpha_3^+ + B_6^+\right)B_2^+}{2B_0^+\alpha_3^+}; \tag{1.92}$$

$$B_7^+ = T_{12} - \frac{\left(3B_4^+\alpha_3^{+2} + 2A_{12}\alpha_3^+ + B_6^+\right)B_3^+}{2B_0^+\alpha_3^+}. \tag{1.93}$$

Substituting the resulting expressions for B_5^+ and B_7^+ into Eq. (1.81) and considering condition (1.59), we have

$$B_8\alpha_3^{+4} + B_9\alpha_3^{+2} + B_{10}\alpha_3^+ + B_{11} = 0, \tag{1.94}$$

where

$$B_8 = -B_0^+B_4^+, \quad B_{10} = 2(T_{12} - H^-)B_0^+ - 2A_{12}B_3^+,$$
$$B_9 = -B_0^+B_6^+ - 3B_3^+B_4^+, \quad B_{11} = -B_3^+B_6^+.$$

In this expression coefficients B_8, B_9, B_{10}, and B_{11} are uniquely defined and the problem thus reduces to the solution of this equation in α_3^+. To solve it, we can use, for instance, the method of tangents or chords [18]. In solving this equations we must find only the values of α_3^+ that satisfy conditions (1.77).

By finding α_3^+ and then the β_1^+ corresponding to it, and by verifying that conditions (1.77) have been satisfied for the values of α_3^+ and β_1^+ thus obtained, let us dwell on the values for which conditions (1.77) are satisfied. If there are no such values, then let us move on to considering the next points in formulas (1.82).

Let us determine for the values α_3^+ and β_1^+ satisfying conditions (1.77) the coefficient N_{12} corresponding to it based on Eq. (1.92). Then, let us substitute the values of a_3^+ into (1.86) and let us define the sign of the function C. If $C < 0$, α_3^+ and β_1^+ are left for further investigation, but if $C > 0$, they are discarded. But if after all this analysis it turns out that there are several solutions to Eq. (1.94) for which all the above-listed conditions are satisfied, then let us dwell on the solution for which the coefficient N_{12} is at a minimum. /50

The value of α^+ and β^+ for the remaining points of formulas (1.82) at the instant of staging are determined directly. By verifying condition (1.77) for these points, we discard the point for which these conditions are not satisfied. For the remaining points we find the sign of function C (see Eq. (186)). For points

2 and 3 we must have $C > 0$, and for points 4, 5, 6, and 7 there must be $C < 0$. By verifying this analysis, we discard the points for which the sign of function C does not correspond to the required sign.

Let us find for the remaining points the coefficient N_{12} corresponding to them, using the formula

$$N_{12}=\frac{B_2^+\left(B_4^+\alpha^{+3}+A_{12}\alpha^{+2}+B_6^+\alpha^++T_{12}-H^-\right)}{B_0^+\alpha^{+2}+B_3^+}. \tag{1.95}$$

Knowing N_{12}, we can find the coefficient B_7^+ of Eq. (1.81). Obviously, if we take as the criterion of the optimal control the control for which the coefficient B_7^+ is the smallest, the function (1.81) will take on the largest value with respect to the control, that is, the control will be optimal. By thus finding the optimal control and the coefficient N_{12} corresponding to it, we can find the multiplier λ_5^+ by formula (1.90).

Solution of the Problem

In the preceding sections we outlined a method for determining the optimal control at a fixed instant of time τ. By thus determining the optimal control, we can now find the multipliers ψ_3, ψ_4, and ψ_5 appearing in this system of equations (1.64).

They are defined from the conditions

$$\left.\begin{aligned}&2A_{1k}\alpha+E_{1k}+F_{1k}\beta=\psi_1(2\alpha-\alpha_{\min}-\alpha_{\max})+\\&+\psi_3(2n_{y1}-n_{y1\min}-n_{y1\max})\frac{c_y^\alpha Sqr^2}{m\mu}+\\&+\psi_4(2A_{2k}\alpha+E_{2k})+\psi_5(2A_{3k}\alpha+E_{3k});\\&F_{1k}\alpha+N_{1k}=\psi_2(2\beta-\beta_{\min}-\beta_{\max})+\psi_4N_{2k}+\psi_5N_{3k}.\end{aligned}\right\} \tag{1.96}$$

Here k = 1 and k = 2 after staging.

As we can easily see, the two equations contain five unknowns. /51
Since there are two control functions, motion is possible only either within the domains (1.46) - (1.50) or else along one or two of the bounds of these domains. Hence it follows that the maximum of the multiplier ψ_j (j = 1, 2, 3, 4, and 5) simultaneously nonzero cannot exceed two.

$\psi_1 \neq 0$, if for the selected optimal control $\xi_1 = 0$;
$\psi_2 \neq 0$, " " " " " " $\xi_2 = 0$;
$\psi_3 \neq 0$, " " " " " " $\xi_3 = 0$;
$\psi_4 \neq 0$, " " " " " " $\xi_4 = 0$;
$\psi_5 \neq 0$, " " " " " " $\xi_5 = 0$.

The multipliers are equal to zero in all the remaining cases.

Next, let us integrate the systems (1.45) and (1.64) with the optimal values of α and β thus obtained, and also ψ_j. In the next stage of integration the entire process for determining the optimal controls α and β, and also the values of ψ_j is repeated over again, and so on.

The conditions for the discontinuity of the multipliers λ_i ($i = 1, 2, 3, 4, 5$) at the instant of insertion are as follows:

-- at the bound $\psi_4 = 0$

$$\left.\begin{aligned} &\lambda_1^+ = \lambda_1^- + \nu\varrho V; \\ &\lambda_2^+ = \lambda_2^-; \\ &\lambda_3^+ = \lambda_3^-; \\ &\lambda_4^+ = \lambda_4^- + \nu\frac{\partial\varrho}{\partial h}\frac{V^2}{2}; \\ &\lambda_5^+ = \lambda_5^-; \end{aligned}\right\} \tag{1.97}$$

-- at the bound $\psi_5 = 0$

$$\left.\begin{aligned} &\lambda_1^+ = \lambda_1^- + \nu_1; \\ &\lambda_2^+ = \lambda_2^-; \\ &\lambda_3^+ = \lambda_3^-; \\ &\lambda_4^+ = \lambda_4^- - \nu_1\frac{\partial l(h)}{\partial h}; \\ &\lambda_5^+ = \lambda_5^-. \end{aligned}\right\} \tag{1.98}$$

Now let us formulate the general solution to the problem.

Since the function H is homogeneous, the initial value of one of the multipliers λ_i can be selected arbitrarily, for example,

$$\lambda_{50} = 1 \tag{1.99}$$

One of the multipliers λ_i, for example, λ_{40}, can be determined from the condition:

$$\lambda_{40} = -\frac{\lambda_{10}f_1 + \lambda_{20}f_2 + \lambda_{30}f_3 + \lambda_{50}f_5}{f_4}. \tag{1.100}$$

When the systems (1.45) and (1.64) are integrated, we have ten arbitrary constants, and the time of completion of insertion t_f is also unknown. /52

To determine the 11 unknowns, we have 11 relations (1.52), (1.53), (1.99) and (.100). Integrating the systems (1.45) and (1.64) under the initial conditions (1.52), (1.99), and (1.100), and using, for example, Newton's method [18] to determine λ_{10}, λ_{20}, and λ_{30}, we satisfy the conditions (1.53).

In solving the problem with constraints as on the first-order phase coordinates, the initial value of one of the multipliers λ_i, for example, λ_{10}, is selected so that condition (1.59) is satisfied at the instant of insertion at the bound $\xi_4 = 0$, and then by selecting the number ν, let us satisfy the conditions (1.59) at the instant of insertion at bound $\xi_5 = 0$ and the number ν_i (equally with λ_{20} and λ_{30}) is determined on the condition that the conditions (1.53) are satisfied.

First, let us dwell on the solution to the boundary value problem without referring to the constraints on the phase coordinates.

Suppose that in the first approximation of the initial multipliers λ_{10}, λ_{20}, and λ_{30} we assumed λ'_{10}, λ'_{20} and λ'_{30}, and that the resulting mismatches of the boundary conditions corresponding to this approximation are as follows:

$$\left.\begin{aligned}\Delta V &= \tilde{V} - V_f,\\ \Delta\theta &= \tilde{\theta} - \theta_f,\\ \Delta L &= \tilde{L} - L_f,\end{aligned}\right|$$

where $\tilde{V}$, $\tilde{\theta}$, and $\tilde{L}$ are the values of V, θ, and L obtained by integrating the systems(1.45) and (1.64) at the instant $\tilde{h} = h_f$.

Varying λ'_{10}, λ'_{20}, and λ'_{30} by a sufficiently small quantity, let us find at the instant $\tilde{h} = h_f$ the partial derivatives

$$\frac{\partial V}{\partial\lambda'_{10}},\quad \frac{\partial V}{\partial\lambda'_{20}},\quad \frac{\partial V}{\partial\lambda'_{30}},\quad \frac{\partial\theta}{\partial\lambda'_{10}},\quad \frac{\partial\theta}{\partial\lambda'_{20}},\quad \frac{\partial\theta}{\partial\lambda'_{30}},\quad \frac{\partial L}{\partial\lambda'_{10}},\quad \frac{\partial L}{\partial\lambda'_{20}},\quad \frac{\partial L}{\partial\lambda'_{30}}.$$

Knowing the partial derivatives, we can compute the correction $\Delta\lambda'_{10}$, $\Delta\lambda'_{20}$, and $\Delta\lambda'_{30}$ to the individual values λ'_{10}, λ'_{20}, and λ'_{30}, for example, by Newton's method, thereby solving the system of three algebraic inhomogeneous equations

$$\left.\begin{aligned}&\frac{\partial V}{\partial\lambda'_{10}}\Delta\lambda'_{10} + \frac{\partial V}{\partial\lambda'_{20}}\Delta\lambda'_{20} + \frac{\partial V}{\partial\lambda'_{30}}\Delta\lambda'_{30} + \varepsilon\Delta V = 0;\\ &\frac{\partial\theta}{\partial\lambda'_{10}}\Delta\lambda'_{10} + \frac{\partial\theta}{\partial\lambda'_{20}}\Delta\lambda'_{20} + \frac{\partial\theta}{\partial\lambda'_{30}}\Delta\lambda'_{30} + \varepsilon\Delta\theta = 0;\\ &\frac{\partial L}{\partial\lambda'_{10}}\Delta\lambda'_{10} + \frac{\partial L}{\partial\lambda'_{20}}\Delta\lambda'_{20} + \frac{\partial\theta}{\partial\lambda'_{30}}\Delta\lambda'_{30} + \varepsilon\Delta L = 0.\end{aligned}\right| \qquad (1.101)$$

The process is repeated until the required accuracy is reached. /53

The quantity ε ($0 < \varepsilon \leq 1$) is selected on the condition that the solution of the boundary value problem converges, where the closer we come to the assigned boundary conditions, the closer ε is taken to 1. The speed with which the optimal solution of this problem is determined depends on the proper choice of ε.

If the first approximation is sufficiently good, we can assume $\varepsilon = 1$. But if it is far from the required value, the value of ε must be taken closer to zero, by increasing its value from one approximation to the other.

If the problem is solved with reference to constraints on the phase coordinate, the process of solving the boundary value problem is somewhat more complicated, since in this case, in addition to everything else, the conditions of discontinuity at the docking moment must be satisfied. In turn, the new multipliers ν and ν_1 will appear, which also must be determined from the solution to the boundary value problem.

To solve this problem, it is useful first to solve the problem without reference to the constraints on the first-order phase coordinates. Then, by assuming to the first approximations the resulting values of λ_{10}, λ_{20}, and λ_{30}, and revising λ_{10}, λ_{20}, λ_{30}, ν, and ν_1, we satisfy all the conditions of discontinuity at the moment of docking and the boundary conditions at the right endpoint of integration.

Let us consider the solution to this problem more closely.

Suppose that we adopt as the first approximation of the initial multipliers λ_{10}, λ_{20}, and λ_{30}, the quantities λ'_{10}, λ'_{20}, and λ'_{30}, and suppose that the function H^+ is not equal to zero at the moment of docking. By varying λ'_{10} by a sufficiently small quantity, at the moment of docking we find the partial derivative $\frac{\partial H^+}{\partial \lambda'_{10}}$. The multipliers λ'_{20} and λ'_{30} are not yet varied.

By knowing the partial derivative, we can compute the correction λ'_{10} for the initial value λ'_{10} based on the formula

$$\Delta\lambda'_{10} = -\varepsilon_1 \frac{H^+}{\partial H^+/\partial\lambda'_{10}}.$$

The process is repeated until the required accuracy in H^+ is attained. The value of ν is revised similarly, only it is carried out now on the condition that the discontinuity is satisfied at the moment of second docking, namely, by assigning the approximate number ν_0 at the instant of first docking, we determine H^+ at the moment of the second docking, therefore varying $\nu_0|$ by a sufficiently small quantity, let us determine the partial derivative $\frac{\partial H^+}{\partial \nu_0}$ at the moment of second docking. Then the correction to the initial value ν_0 will be determined by the formula

$$\Delta\nu_0 = -\varepsilon_2 \frac{H^+}{\partial H^+/\partial \nu_0}.$$

The process is repeated until the required accuracy in H^+ is attained at the moment of second docking. /54

The value of ν_1 is revised on the condition that, when $h = h_f$, condition $V = V_f$ is satisfied, based on the formula

$$\Delta\nu_{10} = -\varepsilon_3 \frac{\Delta V}{\partial V/\partial \nu_{10}}.$$

The trajectory thus selected satisfies all the conditions of discontinuity at the moment of docking and the two boundary conditions (1.53) at the right endpoint of integration.

Then, by varying λ'_{20} and λ'_{30} by a small enough quantity and by solving the analogous problem in the selection of λ_{10}, ν, and ν_1, let us determine the partial derivatives

$$\frac{\partial \theta}{\partial \lambda'_{20}}, \quad \frac{\partial \theta}{\partial \lambda'_{30}}, \quad \frac{\partial L}{\partial \lambda'_{20}}, \quad \frac{\partial L}{\partial \lambda'_{30}}.$$

Knowing the partial derivatives we can compute the corrections $\Delta\lambda'_{20}$ and $\Delta\lambda'_{30}$ to the initial values λ'_{20} and λ'_{30}, by solving the following system of two equations:

$$\frac{\partial \theta}{\partial \lambda'_{20}} \Delta\lambda'_{20} + \frac{\partial \theta}{\partial \lambda'_{30}} \Delta\lambda'_{30} + \varepsilon_4 \Delta\theta = 0;$$
$$\frac{\partial L}{\partial \lambda'_{20}} \Delta\lambda'_{20} + \frac{\partial L}{\partial \lambda'_{30}} \Delta\lambda'_{30} + \varepsilon_4 \Delta L = 0.$$

The process is repeated until the required accuracy in satisfying the boundary conditions (1.53) is attained).

The values of the numbers ε_1, ε_2, ε_3 and ε_4 ($0 < \varepsilon_i \leq 1$, $i = 1, 2, 3, 4$), just as ε, are selected on the condition that the process of solving the problem converges.

Particular Cases of the Overall Problem of Optimizing the Trajectory of LV Insertion into AES Orbit

First let us examine the case when an LPRE is used as the power plant in both the first and second stages. We know that the optimal trajectories of an LV with LPRE are characterized by its rapid departure from the dense atmospheric layers, which enables us to remove the constraints (1.49) and (1.50). Thus, Eqs. (1.56) and (1.57) will be of the form

$$\dot{\lambda}_i = -\frac{\partial H}{\partial x_i} + \psi_3 \frac{\partial \xi_3}{\partial x_i} \quad (i = 1,\ 2,\ 3,\ 4,\ 5); \tag{1.102}$$

$$\frac{\partial H}{\partial u} = \psi_1 \frac{\partial \xi_1}{\partial u} + \psi_2 \frac{\partial \xi_2}{\partial u} + \psi_3 \frac{\partial \xi_3}{\partial u} \quad (u = \alpha,\ \beta). \tag{1.103}$$

Condition (1.59) in this case (with the exception of the moment of staging) will be automatically satisfied, and conditions (1.60) and (1.61) will not be used. /55

Eqs. (1.64) can be written thusly

$$\left.\begin{aligned}
\dot{\lambda}_1 &= -\lambda_1 \frac{\partial f_1}{\partial V} - \lambda_2 \frac{\partial f_2}{\partial V} - \lambda_3 \frac{\partial f_3}{\partial V} - \lambda_4 \frac{\partial f_4}{\partial V} + \psi_3 \frac{\partial \xi_3}{\partial V};\\
\dot{\lambda}_2 &= -\lambda_1 \frac{\partial f_1}{\partial \theta} - \lambda_2 \frac{\partial f_2}{\partial \theta} - \lambda_3 \frac{\partial f_3}{\partial \theta} - \lambda_4 \frac{\partial f_4}{\partial \theta};\\
\dot{\lambda}_3 &= 0;\\
\dot{\lambda}_4 &= -\lambda_1 \frac{\partial f_1}{\partial h} - \lambda_2 \frac{\partial f_2}{\partial h} - \lambda_3 \frac{\partial f_3}{\partial h} + \psi_3 \frac{\partial \xi_3}{\partial h};\\
\dot{\lambda}_5 &= -\lambda_1 \frac{\partial f_1}{\partial m} - \lambda_2 \frac{\partial f_2}{\partial m} + \psi_3 \frac{\partial \xi_3}{\partial m}.
\end{aligned}\right\} \tag{1.104}$$

The partial derivatives $\frac{\partial f_1}{\partial V}$, $\frac{\partial f_2}{\partial V}$, $\frac{\partial f_3}{\partial V}$, $\frac{\partial f_4}{\partial V}$, $\frac{\partial \xi_3}{\partial V}$, $\frac{\partial f_1}{\partial \theta}$, $\frac{\partial f_2}{\partial \theta}$, $\frac{\partial f_3}{\partial \theta}$, $\frac{\partial f_4}{\partial \theta}$, $\frac{\partial f_1}{\partial h}$, $\frac{\partial f_2}{\partial h}$, $\frac{\partial f_3}{\partial h}$, $\frac{\partial \xi_3}{\partial h}$, $\frac{\partial f_1}{\partial m}$, $\frac{\partial f_2}{\partial m}$, $\frac{\partial \xi_3}{\partial m}$ are determined from the same formulas as in Eqs. (1.64). The engine thrust in both the first and second stages will be determined by Eq. (1.67), and its partial derivatives will be of the form

$$\frac{\partial P}{\partial V} = 0, \quad \frac{\partial P}{\partial h} = -g_0 e_0 S_a \frac{\partial \varrho}{\partial h}. \tag{1.105}$$

But the formulas (1.65) and (1.66) are not used.

The coefficient A_{11}, E_{11}, F_{11}, F_{11}, N_{11} and T_{11} of Eq. (1.70) will be of the form of the corresponding coefficients A_{12}, E_{12},

F_{12}, N_{12}, and T_{12} of Eq. (1.71). Their only difference will be that in the computation of coefficients with subscripts "11" we must take all the characteristics of the LV and the power plant corresponding to the first stage, while when computing the coefficients with subscripts "12" we must take the corresponding characteristics of the LV and the power plant of the second stage.

Because the constraints (1.49) and (1.50) are removed, the case of considering the motion of an LV along the bounds ξ_4 and ξ_5 disappears and, naturally, the formulas (1.72), (1.73), (1.74), (1.75), (1.80), (1.81), (1.82), (1.91), (1.92), (1.93), (1.94), (1.95), (1.97), and (1.98), just as the solution of the boundary value problem with reference to the constraints on the first-order phase coordinates, will become unnecessary.

In the equations (1.96), considering the foregoing, we must adopt $\psi_4 = \psi_5 = 0$. All the remaining formulas, which are not mentioned here, are used in the form in which they are represented in the method given earlier.

In conclusion, let us dwell on the insertion of an LV into an AES orbit, beginning with the end of the atmospheric section of the flight when LPRE are used in the second and third stages, and adopting as the specific thrust the specific thrust of the LPRE in vacuo. /56

In this case, system (1.45) becomes:

$$\left.\begin{aligned}
&\dot V=\frac{J_0 g_0\beta}{m}-\frac{\mu}{r^2}\sin\theta=f_1;\\
&\dot\theta=\frac{J_0 g_0\beta\alpha}{mV}-\frac{\mu}{r^2V}\cos\theta+\frac{V\cos\theta}{r}=f_2;\\
&\dot L=\frac{r_0}{r}V\cos\theta=f_3;\\
&\dot h=V\sin\theta=f_4;\\
&\dot m=-\beta=f_5.
\end{aligned}\right\}\qquad(1.106)$$

Naturally, in this formulation of the problem we must consider only the constraints (1.46) and (1.47), while we must discard the constraints (1.48), (1.49) and (1.50).

The system of conjugate equations (1.56) in this case will be of the form:

$$\dot\lambda_i=-\frac{\partial H}{\partial x_i}\,(i=1,\ 2,\ 3,\ 4,\ 5).\qquad(1.107)$$

Let us write it in expanded form:

$$\left.\begin{aligned}
\dot{\lambda}_1 &= \lambda_2\left(\frac{J_0 g_0 \beta\alpha}{mV^2} - \frac{\mu}{r^2V^2}\cos\theta - \frac{\cos\theta}{r}\right) - \lambda_3\frac{r_0}{r}\cos\theta - \lambda_4\sin\theta;\\
\dot{\lambda}_2 &= \lambda_1\frac{\mu}{r^2}\cos\theta - \lambda_2\left(\frac{\mu}{r^2V} - \frac{V}{r}\right)\sin\theta + \lambda_3\frac{r_0}{r}V\sin\theta - \lambda_4 V\cos\theta;\\
\dot{\lambda}_3 &= 0;\\
\dot{\lambda}_4 &= -\lambda_1\frac{2\mu}{r^3}\sin\theta - \lambda_2\left(\frac{2\mu}{r^2V} - \frac{V}{r^2}\right)\cos\theta + \lambda_3\frac{r_0}{r^2}V\cos\theta;\\
\dot{\lambda}_5 &= \frac{J_0 g_0 \beta}{m^2}\left(\lambda_1 + \lambda_2\frac{\alpha}{V}\right).
\end{aligned}\right\} \qquad (1.108)$$

The function H both before separation as well as after separation will be of the form

$$H = F\alpha\beta + N\beta + T, \qquad (1.109)$$

where

$$\left.\begin{aligned}
F &= \lambda_2\frac{J_0 g_0}{mV};\\
N &= \lambda_1\frac{J_0 g_0}{m} - \lambda_5;\\
T &= -\lambda_1\frac{\mu}{r^2}\sin\theta + \lambda_2\left(\frac{V}{r} - \frac{\mu}{r^2V}\right)\cos\theta +\\
&\quad + \lambda_3\frac{r_0}{r}V\cos\theta + \lambda_4 V\sin\theta.
\end{aligned}\right\}$$ /57

Since $\frac{J_0 g_0}{mV} > 0$, the sign of coefficient F is uniquely determined by the sign of multiplier λ_2. Since $\beta \geq 0$, it is clear that the sign of $F\beta$ will be characterized by the sign of λ_2. And since α appears linearly in the function H, the optimal control in terms of the angle of attack α can take on only limiting values, where if $\lambda_2 < 0$, we must taken $\alpha_{opt} = \alpha_{min}$; if $\lambda_2 > 0$, $\alpha_{opt} = \alpha_{max}$.

Similarly, we can show that only the limiting values defined as follows can be optimal controls in terms of the mass flow rate of fuel β: $\beta_{opt} = \beta_{min}$, if $F\alpha + N < 0$; $\beta_{opt} = \beta_{max}$, if $F\alpha + N > 0$.

Thus, the optimal control in any fixed moment of time τ (except for the staging moment) is uniquely defined.

Now let us dwell on selecting the optimal control at the instant of staging. In the case examined here, the problem of determining the optimal control at the instant of staging is solved quite readily.

Actually, at the instant of staging only the multiplier λ_5 suffers a discontinuity (see Eq. (1.83)). This multiplier appears in the function H in terms of multiplier N. Since the angle of attack α is entirely defined by the sign of multiplier λ_2 (and this multiplier at the instant of staging is continuous), it is clear that the angle of attack α^+ at the instant of staging is determined uniquely based on the method indicated above.

By computing the coefficient T^+ (we know that $T^+ = T^-$, since the rocket parameters do not appear in the multiplier T), we find its sign. $T^+ < 0$, then because $H^+ = 0$, it follows that $(F^+\alpha^+ + N^+)\beta^+ > 0$, or expressed another way:

$$F^+\alpha^+ + N^+ > 0.$$

And this means that we must take $\beta^+ = \beta_{max}$ as the optimal control at the instant of staging.

But if $T^+ > 0$, we must take $\beta^+ = \beta_{min}$ as the optimal control.

The coefficients N^+ and λ_5^+ corresponding to the control thus selected are determined by the formulas

$$N^+ = -\frac{T^+ + F^+\alpha^+}{\beta^+};$$

$$\lambda_5^+ = \lambda_1^- \frac{J_0^+ g_0}{m^+} - N^+.$$

1.3. Method of Inserting Satellite into Orbit Without Final Combustion of Fuel /58

Let us examine the case when a satellite is inserted into orbit with a continuous powered trajectory section of the launch vehicle. Obviously, the shape of the orbit and its distance at different points from the Earth's surface will depend on the selection of the elements of the end of the powered section, and also on the adopted law of launch vehicle motion. For many kinds of satellites it is important that their orbital perigee lie as high as possible. Accordingly, we will seek the maximum altitude of the end of the powered section for an assigned final velocity and assigned direction of this velocity.

We will not dwell on methods of calculating the powered section itself, assuming it to be known to the reader from already published studies (for example, [5]).

Let us determine the optimal law of launch vehicle motion along the powered section, that is, let us find the optimal program of variation in its pitch angle ϕ_{pr}, without referring to the variable fuel consumption. The methods examined below are a further elaboration of the study [46]. They permit the quite the quite simple selection of ϕ_{pr} virtually without loss of accuracy. The optimal ϕ_{pr} found can be used in an exact calculation of the powered section for a specific launch vehicle.

Optimal Program for Angle of Pitch with Constant Fuel Consumption over the Powered Section

We will start from the system of differential equations of motion that take into account the curvature of the Earth and the variation in gravity with altitude [46]. For the case of constant fuel consumption this enables us to obtain a simple analytic method of calculation:

$$\left.\begin{aligned} \frac{du}{dt} &= p\cos\varphi - \nu^2 x; \\ \frac{dw}{dt} &= p\sin\varphi + 2\nu^2 y - g_0; \\ \frac{dx}{dt} &= u; \\ \frac{dy}{dt} &= w, \end{aligned}\right| \tag{1.110}$$

where

$$\nu^2 = \frac{g_0}{R};$$

g_0 is the acceleration of the force of gravity at zero altitude; R is the Earth's radius; p is launch vehicle acceleration; and ϕ is the angle of pitch.

The system of equations (1.110) can be readily obtained from the ordinary equations of rocket plane motion written in an absolute rectangular system of coordinates with the origin at the launch point, the Oy axis directed upward along the vertical, and the Ox axis directed toward the powered section of the trajectory. /59

By expanding in power series the expressions for the projections of gravitational accelerations

$$g_x = g_0 \frac{R^2 x}{[x^2 + (y+R)^2]^{3/2}};$$
$$g_y = g_0 \frac{R^2 (y+R)}{[x^2 + (y+R)^2]^{3/2}}$$

and by limiting ourselves to the first terms in the expansion, we get

$$g_x = g_0 \frac{x}{R};$$
$$g_y = g_0 - 2g_0 \frac{y}{R}.$$

In the particular case when we are considering a two-stage rocket, we will take as the initial point the end of the powered section of the first stage, designating it by the subscript "1".

Let us set up the variational problem of finding the function $\phi(t)$ for assigned boundary conditions that give a maximum to the altitude

$$h_f = y_f + \frac{x_f^2}{2R}, \tag{1.11}$$

where the subscript "f" refers to the elements of the end of the powered section.

It should be noted that this problem is equivalent to the problem of determining the function $\phi(t)$ that gives a maximum for the velocity at assigned altitude [46]. From the standpoint of the calculus of variations, this is a problem of the conditional extremum with differential relations. The equations of motion (1.110) of the launch vehicle are the relations.

Let us set up the Lagrangian function

$$H = \lambda_1\psi_1 + \lambda_2\psi_2 + \lambda_3\psi_3 + \lambda_4\psi_4,$$

where λ_1, λ_2, λ_3, and λ_4 are certain differentiable functions with respect to time t:

$$
\begin{aligned}
\psi_1 &= \frac{du}{dt} + \nu^2 x - p\cos\varphi;\\
\psi_2 &= \frac{dw}{dt} - 2\nu^2 y - p\sin\varphi + g_0;\\
\psi_3 &= \frac{dx}{dt} - u;\\
\psi_4 &= \frac{dy}{dt} - w.
\end{aligned}
$$

The functional whose maximum we seek is of the form /60

$$
I = h_f + \int_0^{t_f} H\,dt. \tag{1.112}
$$

Let us write the first variation for I:

$$
\delta I = \delta h_f + \int_0^{t_f} \delta H\,dt, \tag{1.113}
$$

where

$$
\begin{aligned}
\delta h &= \delta y_f + \frac{x}{R}\delta x;\\
\delta H &= H'_u\delta u + H'_w\delta w + H'_y\delta y + H'_x\delta x + H'_{\dot u}\delta\dot u +\\
&\quad + H'_{\dot w}\delta\dot w + H'_{\dot y}\delta\dot y + H'_{\dot x}\delta\dot x + H'_{\varphi}\delta\varphi.
\end{aligned}
$$

Canceling out the derivatives $\delta\dot{u}$, $\delta\dot{w}$, $\delta\dot{y}$, and $\delta\dot{x}$ by integrating the corresponding expressions by parts, we get

$$
\begin{aligned}
\delta I &= \delta y_f + \frac{x_f}{R}\delta x_f + [\lambda_1\delta u]_0^{t_f} + [\lambda_2\delta w]_0^{t_f} + [\lambda_3\delta x]_0^{t_f} +\\
&+ [\lambda_4\delta y]_0^{t_f} + \int_0^{t_f}[(-\lambda_3 - \dot\lambda_1)\delta u - (\lambda_4 + \dot\lambda_2)\delta w + (\lambda_1\nu^2 - \dot\lambda_3)\delta x -\\
&- (2\nu^2\lambda^2 + \dot\lambda_4)\delta y + p(\lambda_1\sin\varphi - \lambda_2\cos\varphi)\delta\varphi]\,dt.
\end{aligned} \tag{1.114}
$$

The extremal value of the functional I corresponds to the zero-equality of its first variation δI.

By setting equal to zero the multipliers appearing under the sign of the integral in front of the variations δu, δw, δx, and δy, we get a system of differential equations for the quantities λ_1, λ_2, λ_3, and λ_4 and the final relation for ϕ:

$$\begin{aligned} &\dot{\lambda}_1 + \lambda_3 = 0; \\ &\dot{\lambda}_2 + \lambda_4 = 0; \\ &\dot{\lambda}_3 - \nu^2\lambda_2 = 0; \\ &\dot{\lambda}_4 + 2\nu^2\lambda_2 = 0; \\ &\lambda_1 \sin\varphi - \lambda_2 \cos\varphi = 0. \end{aligned} \qquad (1.115)$$

From the last equation we have

$$\operatorname{tg}\varphi = \frac{\lambda_2}{\lambda_1}. \qquad (1.116)$$

The solution to the system of equations (1.115) is of the form: /61

$$\begin{aligned} &\lambda_1 = C_1 \cos \nu t + C_2 \sin \nu t; \\ &\lambda_2 = C_3 \operatorname{sh} \sqrt{2}\nu t + C_4 \operatorname{ch} \sqrt{2}\nu t; \\ &\lambda_3 = C_1 \nu \sin \nu t - C_2 \nu \cos \nu t; \\ &\lambda_4 = -C_3 \sqrt{2}\nu \operatorname{ch} \sqrt{2}\nu t - C_4 \sqrt{2}\nu \operatorname{sh} \sqrt{2}\nu t. \end{aligned} \qquad (1.117)$$

Eqs. (1.115) and formula (1.116) are analogous to the corresponding expressions in [46]. Let us select the following boundary conditions. At the end of the powered section we will fix the velocity vector (u_f, w_f), and at the initial point -- the modulus of velocity, that is, at the end of the powered section of the first stage the following relation must be satisfied

$$v_1^2 = u_1^2 + w_1^2 = \text{const}, \qquad (1.118)$$

which gives for the variations δu_1 and δw_1 the relation

$$\delta u_1 = -\frac{w_1}{u_1}\,\delta w_1. \qquad (1.119)$$

We will also fix the coordinates x and y of the initial point.

This selection of the boundary conditions simplifies the solution of the problem. It is quite justified, since assigning a constant velocity modulus at the initial point instead of the functional dependence of velocity on altitude and angle of inclination of the orbit is based on compensating for the effect of change in velocity at the initial point on the insertion altitude h_f by the influence of the corresonding change in altitude (in angle) at this same point.

It is convenient to take as the initial point the end of the powered section of the first stage, since in the first stage of flight occurring in the dense atmospheric layers, ϕ_{pr} is selected on the condition that the angles of attack are near-zero, and not on the condition that the maximum altitude of the insertion point is obtained.

Let us examine the case of a two-stage launch vehicle. As the initial point we will take the end of the powered section of the first stage. We will assume that the program for this section is assigned, more exactly, that is depends on one parameters -- the maximum angle of attack α_m. By selecting α_m the direction of the tangent to the trajectory at the beginning of the flight of the second stage is determined, that is, the angle θ_1.

Depending on the values of the angle of the velocity vector θ_{extra} obtained by solving the variational problem in each specific case, we will rotate the entire first section of the trajectory so that the angle θ_1 coincides with $\theta_{1\,extra}$ and then we will repeat the calculation with the revised values of v_1, x_1, and y_1. It can be assumed that this method gives a better approximation to the extremal program for the entire powered section as a whole.

By substituting δu_1 from Eq. (1.119) into the expression (1.114) and considering that the variations δu_f and δy_f at the end of the powered section are equal to zero, and δx_f and δy_f are arbitrary, we get /62

$$\lambda_{11}\frac{w_1}{u_1}-\lambda_{21}=0.$$

Since $w_1 = v_1 \sin\theta_1$ and $u_1 = v_1 \cos\theta_1$,

$$\operatorname{tg}\theta_1=\frac{\lambda_{21}}{\lambda_{11}}=\operatorname{tg}\varphi_1; \tag{1.120}$$

$$\left.\begin{aligned}\lambda_{4f}&=-1;\\ \lambda_{3f}&=-\frac{x_f}{R}.\end{aligned}\right\} \tag{1.121}$$

On the other hand, from system (1.118) it follows that when t = 0, we have

$$\left.\begin{aligned}\lambda_{11}&=C_1;\\ \lambda_{21}&=C_4.\end{aligned}\right\} \tag{1.122}$$

where $t = t_f$

$$\left.\begin{aligned}\lambda_{1f} &= C_1 \cos \nu t_f + C_2 \sin \nu t_f;\\ \lambda_{2f} &= C_3 \operatorname{sh} \sqrt{2}\nu t_f + C_4 \operatorname{ch} \sqrt{2}\nu t_f;\\ \lambda_{3f} &= C_1 \nu \sin \nu t_f - C_2 \nu \cos \nu t_f;\\ \lambda_{4f} &= -C_3 \sqrt{2}\nu \operatorname{ch} \sqrt{2}\nu t_f - C_4 \sqrt{2}\nu \operatorname{sh} \sqrt{2}\nu t_f.\end{aligned}\right\} \qquad (1.123)$$

From the conditions (1.121) and (1.23) we obtain two equations for the four unknown coefficients C_1, C_2, C_3, and C_4:

$$\left.\begin{aligned}C_1 \nu \sin \nu t_f - C_2 \nu \cos \nu t_f &= -\frac{x_f}{R};\\ C_3 \sqrt{2}\nu \operatorname{ch} \sqrt{2}\nu t_f + C_4 \sqrt{2}\nu \operatorname{sh} \sqrt{2}\nu t_f &= 1\end{aligned}\right\} \qquad (1.123')$$

or, by solving (1.123'), for C_2 and C_3, we have

$$C_2 = \frac{\frac{x_f}{R} + C_1 \nu \sin \nu t_f}{\nu \cos \nu t_f};$$

$$C_3 = \frac{1 - C_4 \sqrt{2}\nu \operatorname{sh} \sqrt{2}\nu t_f}{\sqrt{2}\nu \operatorname{ch} \sqrt{2}\nu t_f}.$$

We will find the two equations we lack by using the system of equations of motion (1.110). From the first two we get

$$\frac{w_1}{u_1} = \operatorname{tg} \theta_1 = \frac{\int_0^{t_f} p \sin \varphi dt - g_0 t_f - w_f + \Delta w}{\int_0^{t_f} p \cos \varphi dt - u_f + \Delta u}, \qquad (1.124)$$

where

$$\left.\begin{aligned}\Delta w &= 2\nu^2 \int_0^{t_f} y dt\\ \Delta u &= -\nu^2 \int_0^{t_f} x dt\end{aligned}\right\}$$ are the corrections that allow for the nonparallelity of the field of gravity and the change in acceleration with altitude; (1.125)

$$\varphi=\operatorname{arctg}\frac{\lambda_2}{\lambda_1}=\operatorname{arctg}\frac{C_3\operatorname{sh}\sqrt{2}\nu t+C_4\operatorname{ch}\sqrt{2}\nu t}{C_1\cos\nu t+C_2\sin\nu t}. \tag{1.126}$$

The quantities w_f and u_f in (1.124) are assigned.

Considering that Δw and Δu are small and act as corrections, to the first approximation we can neglect them. Then (1.124) can be considered as one of the two equations we lacked for determining the arbitrary constants C_1, C_2, C_3, and C_4 (or instead of C_4, we can take tg θ_1, since $C_4 = C_1$ tg θ_1). The fourth and last equation is condition (1.118), in which u_1 and w_1 are the numerator and denominator of (1.124). Thus, to the first approximation in solving the problem of selecting the program $\phi(t)$ that gives the greatest altitude for the powered section, with assigned values of the projections of velocity u_f and w_f, we must determine the quantities C_1, C_2, C_3 and tg θ_1, by using Eqs. (1.123), (1.124) and (1.118). To do this, we can assign several values of C_1 (in terms of which all the remaining unknowns C_2, C_3, and C_4 are expressed) and for each of them find from (1.124) by selection the θ_1 corresponding to it. By then substituting the pairs of C_1 and θ_1 we have found into Eq. (1.118), from the series of resulting v_1 we select the value corresponding to the assigned value. Next, we compute the approximate values of the coordinates $\tilde{x}$ and $\tilde{y}$ based on the formula

$$\left.\begin{aligned}\tilde{x}&=x_1+u_1t+\int_0^t d\eta\int_0^\eta p\cos\varphi\, d\xi;\\ \tilde{y}&=y_1+w_1t+\int_0^t d\eta\int_0^\eta (p\sin\varphi-g_0)\, d\xi.\end{aligned}\right\} \tag{1.127}$$

Let us use the resulting values $\tilde{x}(t)$ and $\tilde{y}(t)$ for computing the corrections Δu and Δw with which the entire problem is now solved in the second approximation. Simultaneously the assigned v_1, x_1, and y_1 are revised, as stated above. In addition, by using x_f and y_f, we can, if desired, obtain more exactly the assigned direction of the velocity vector at the end of the power section (for assigned $\theta_f = 0$), and revise w_f using the formula

$$w_f=-\frac{\tilde{x}_f}{R+\tilde{y}_f}u_f. \tag{1.128}$$

Two approximations are sufficient for the solution to the problem. The final values of the coordinates are calculated by the formulas /64

$$\left.\begin{aligned} x&=x_1'+u_1't+\int_0^t d\eta\int_0^\eta p\cos\varphi d\xi+\Delta x;\\ y&=y_1'+w_1't+\int_0^t d\eta\int_0^\eta p\sin\varphi d\xi-\frac{g_0t^2}{2}+\Delta y.\end{aligned}\right\} \tag{1.129}$$

Here x'_1, y'_1, u'_1, and w'_1 are the following values corrected through the first (or second, if required) solution to the problem:

$$\left.\begin{aligned} \Delta x&=-\int_0^t d\eta\int_0^\eta \nu^2 x d\xi;\\ \Delta y&=\int_0^t d\eta\int_0^\eta 2\nu^2 y d\xi.\end{aligned}\right\} \tag{1.130}$$

To verify these working results, it is useful to integrate the equations of motion (1.110) with the above-obtained values of $\phi_{pr}(t)$.

We can considerably simplify the problem in practice without reducing accuracy if the program $\phi_{pr}(t)$ is found for a plane Earth and a parallel field of gravitational force, and if in the equations of motion (1.110), we take into account all the additional terms, regarding them as corrections -- certain functions of time.

For the plane Earth, instead of (1.24), we get

$$\frac{w_1}{u_1}=\operatorname{tg}\theta_1=\frac{\int_0^{t_f} p\sin\operatorname{arctg}\left(\frac{t}{C_2}+\operatorname{tg}\theta_1\right)dt-g_0t_f-w_f+\Delta w}{\int_0^{t_f} p\cos\operatorname{arctg}\left(\frac{t}{C_2}+\operatorname{tg}\theta_1\right)dt-u_f+\Delta u}, \tag{1.131}$$

where C_2 is selected so that Eq. (1.118) is satisfied and Δu and Δw are determined by (1.125), as in the first variant of the problem.

If it is assumed that the accleration of the launch vehicle $p = \frac{n_0 g_0}{1 - \beta t}$, where n_0 is the initial g-load, and β is the relative

fuel consumption ($\beta = \dot{G}/G_0$), the integrals $\int_0^{t_f} p \sin \phi dt$ and /65
$\int_0^{t_f} p \cos \phi dt$ can be written as

$$\left.\begin{aligned}
&\int_0^{t_f} p \sin \operatorname{arctg}\left(\frac{t}{C_2}+\operatorname{tg}\theta_1\right) dt = -\frac{n_0 g_0 C_2}{\bar{b}^2}\left[\frac{1}{a^{1/2}} \ln\left|2(aX)^{1/2}+\right.\right. \\
&\left.\left.+2ax+b\right| + \frac{\bar{a}}{c^{1/2}} \ln\left|\frac{2(cX)^{1/2}}{x}+\frac{2c}{x}+b\right|\right]_0^{t_f}; \\
&\int_0^{t_f} p \cos \operatorname{arctg}\left(\frac{t}{C_2}+\operatorname{tg}\theta_1\right) dt = -\frac{n_0 g_0 C_2}{\bar{b} c^{1/2}} \ln\left|\frac{2(cX)^{1/2}}{x}+\frac{2c}{x}+b\right|_0^{t_f}.
\end{aligned}\right\} \quad (1.32)$$

Here

$$X = ax^2 + bx + c;$$

$$\bar{a} = 1 + \beta C_2 \operatorname{tg}\theta_1; \quad \bar{b} = -\beta C_2; \quad x = \bar{a} + \bar{b}z; \quad z = \frac{t}{C_2} + \operatorname{tg}\theta_1;$$

$$a = 1/\bar{b}^2; \quad b = -2\bar{a}/\bar{b}; \quad c = 1 + \bar{a}^2/\bar{b}^2.$$

Based on these formulas, calculations were made for one of the possible variants of the two-stage launch vehicle for which the dependence of the elements of the powered trajectory section at the moment of staging on the angle of the velocity vector, in turn dependent on the angle of attack α_m, was determined in advance.

Figs. 1.7-1.9 give the altitudes at the end of the powered section (the orbital perigee h_π) as a function of velocity v_f for its horizontal direction for three values of the relative satellite weight G_{sa}/G_0 = 0.027, 0.036, and 0.045 (G_0 is launch weight). In the same three figures are presented the curves of the apogee altitude h_α (the point most distant from the Earth) and the approximate number of satellite revolutions around the Earth during its lifetime.

In the calculation of the insertion section it was assumed that $\beta = 0.005$ and $n_0 = 1.5$ for the second stage. The reduction in the payload in the corresponding variants was compensated by a proportional increase in the fuel weight, and, therefore, in the flight time.

The curves were plotted in terms of the magnitude of velocity at the end of the powered trajectory section v_f which the rocket attained.

Analysis of the graphs leads to the conclusion that there is an optimal number of satellite orbits (that is, its lifetime) which

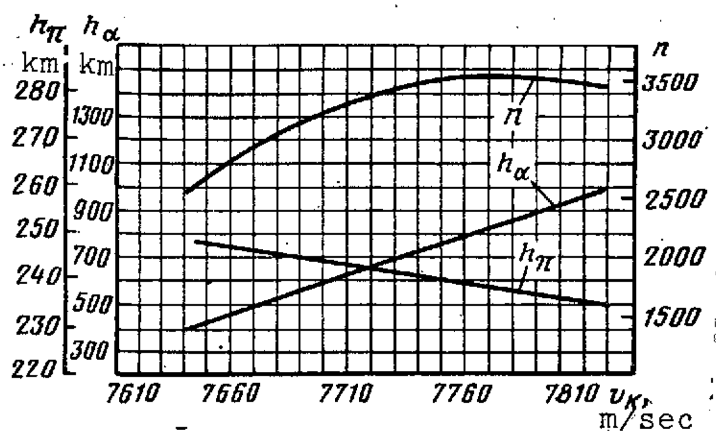


Fig. 1.7. Dependence of h_π, h_α, and n on v_f (G_{sa}/G_0 = 0.027).

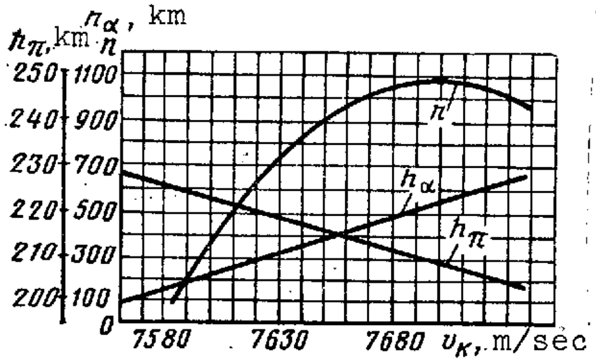


Fig. 1.8. Dependence of h_π, h_α, and n on v_f (G_{sa}/G_0 = = 0.036).

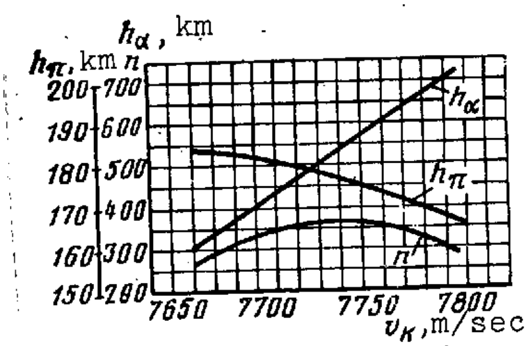


Fig. 1.9. Dependence of h_π, h_α, and n on v_f (G_{sa}/G_0 = 0.045).

corresponds to wholly specific values h_π and h_α in each case. For the variants under consideration, the optimal orbits are quite elongated ellipses, and the oblongness increases with decreasing satellite weight.

These calculations /67
of the trajectories of insertion into orbit where β = const showed that finding the optimal insertion program is a laborious process. A particularly large amount of time was expended in computing by the trial and error method the integrals appearing in Eqs. (1.131) and obtaining the required corrections. On the other hand, selection of the program and calculations of the insertion trajectory can be done sufficiently accurately, since here the optimal characteristics of the orbit are determined or, if the orbit is assigned, the maximum satellite weight.

To accelerate the calculation of the optimal insertion programs, we will assume the initial equations to be in the dimensionless form.

The initial system of equations and the formulas for calculating the optimal program (two-stage launch vehicle) is as follows:

1.
$$\operatorname{tg}\theta_0 = \tau_0 = \frac{w_0}{u_0} = \frac{\int_0^{t_f} p \sin \operatorname{arctg}(bt+\tau_0)\,dt - g_0 t_f - w_f + \Delta w}{\int_0^{t_f} p \cos \operatorname{arctg}(bt+\tau_0)\,dt - u_f + \Delta u}; \qquad (1.133)$$

2.

$$\left.\begin{aligned} v_0^2 &= w_0^2 + u_0^2; \\ \Delta w &= 2\nu^2 \int_0^{t_f} y\,dt; \\ \Delta u &= -\nu^2 \int_0^{t_f} x\,dt; \end{aligned}\right\} \qquad (1.134)$$

3.

$$\tilde{x}_f = x_0 + u_0 t_f + \int_0^{t_f} d\xi \int_0^{\xi} p \cos \operatorname{arctg}(b\eta + \tau_0)\,d\eta;$$

$$\tilde{y} = y_0 + w_0 t_f + \int_0^{t_f} d\xi \int_0^{\xi} p \sin \operatorname{arctg}(b\eta + \tau_0)\,d\eta - g_0 \frac{t_f^2}{2};$$

4.

$$x_f = x_0' + u_0' t_f + \int_0^{t_f} d\xi \int_0^{\xi} p \cos \operatorname{arctg}(b\eta + \tau_0)\,d\eta + \Delta x; \qquad \Delta y = 2\nu^2 \int_0^{t_f} d\xi \int_0^{\xi} y\,d\eta;$$

$$y_f = y_0' + w_0' t_f + \int_0^{t_f} d\xi \int_0^{\xi} p \sin \operatorname{arctg}(b\eta + \tau_0)\,d\eta - g_0 \frac{t_f^2}{2} + \Delta y; \qquad \Delta x = -\nu^2 \int_0^{t_f} d\xi \int_0^{\xi} x\,d\eta;$$ /68

5. $w_f = -u_f \dfrac{x_f}{R + y_f}$ (for the case of insertion when $\theta_f = 0$);

6. $h_f = y_f + \dfrac{x_f}{2R}$.

The following notation was adopted in these equations:

$p = \dfrac{n_0 g_0}{1 - \beta t}$ is the launch vehicle acceleration

n_0 is the initial g-load

β is the fuel consumption coefficient

b, τ_0 are parameters determining the program

$\tau_0 = \operatorname{tg}\theta_0$ is the slope of the initial velocity vector with respect to the Ox axis

$$\nu^2 = \frac{g_0}{R};$$

Δw, Δu are the additional terms that allow for the effect on velocity of the nonparallelity of the gravitational force field and the variation in g with altitude

$\tilde{x}$, $\tilde{y}$ are the coordinates of the launch vehicle calculated without reference to corrections x and y that allow for the effect of the same parameters as for w and y

x_0, y_0, w_0, u_0 are the elements of the beginning of the trajectory section for which the optimal program is determined; the same quantities with the stroke refer to the field of revision

x_f, y_f, w_f, u_f are the elements of the end of the powered section, and

h_f is the altitude of the end of the power section.

The process of calculation reduces to the solution of Eqs. (1.133) and (1.34) for the unknown b and τ_0, with given x_0, y_0, v_0, w_f, and u_f.

In the first approximation the equations are solved without reference to the correction Δw and Δu. Based on the values of u_0, w_0, b, and τ_0 found, let us find the approximate values of the coordinates x and y on the basis of which the corrections Δw, Δu, Δu, and Δx are calculated. In addition, from previously prepared graphs for the functions x_0, y_0, and v_0 on θ_0, characteristic of this launch vehicle, let us compute the new revised quantities x_0', y_0', and v_0' based on the $\theta_0(\tau)$ value found. With these new values of the elements of the initial point the calculation is repeated, that is, we again select b, τ_0, u_0, and w_0. We note that if the initially assigned values prove to be unsuccessful, it is best (if we intend to obtain the horizontal direction of the velocity vector v_f) to revise w_f by means of Eq. (1.128). The process of approximation can be repeated several times, depending on the required accuracy. /69

Finally, after carrying out all the revisions the altitude of the powered section h_f is computed.

Let us bring the integrals appearing in Eq. (1.133) to the dimensionless form. To do this, we introduce the dimensionless time $\xi = t/t_f$.

The integrals become:

$$\left.\begin{aligned} &\int_0^{t_f} p \sin \operatorname{arctg}(bt+\tau_0)\,dt = n_0 g_0 t_f I_1, \\ &I_1 = \int_0^1 \frac{\sin \operatorname{arctg}(b'\xi+\tau_0)}{1-a\xi}\,d\xi; \end{aligned}\right\} \qquad (1.35)$$

$$\left.\begin{aligned}&\int_0^{t_f} p\cos\operatorname{arctg}(bt+\tau_0)\,dt=n_0g_0t_fI_2,\\&I_2=\int_0^1\frac{\cos\operatorname{arctg}(b'\xi+\tau_0)}{1-a\xi}\,d\xi;\end{aligned}\right\}\tag{1.36}$$

$$\left.\begin{aligned}&\int_0^{t_f} d\eta\int_0^{\eta} p\sin\operatorname{arctg}(bt+\tau_0)\,dt=n_0g_0t_f^2I_3,\\&I_3=\int_0^1 d\eta\int_0^{\eta}\frac{\sin\operatorname{arctg}(b'\xi+\tau_0)}{1-a\xi}\,d\xi;\end{aligned}\right\}\tag{1.37}$$

$$\left.\begin{aligned}&\int_0^{t_f} d\eta\int_0^{\eta} p\cos\operatorname{arctg}(bt+\tau_0)\,dt=n_0g_0t_f^2I_4,\\&I_4=\int_0^1 d\eta\int_0^{\eta}\frac{\cos\operatorname{arctg}(b'\xi+\tau_0)}{1-a\xi}\,d\xi;\end{aligned}\right\}\tag{1.38}$$

$$\left.\begin{aligned}&\int_0^{t_f} y\,dt=n_0g_0t_f^3I_5,\\&I_5=\int_0^1 d\xi\int_0^{\xi} d\eta\int_0^{\eta}\frac{\sin\operatorname{arctg}(b'\xi+\tau_0)}{1-a\xi}\,d\xi;\end{aligned}\right\}\tag{1.39}$$

/70

$$\left.\begin{aligned}&\int_0^{t_f} x\,dt=n_0g_0t_f^3I_6,\\&I_6=\int_0^1 d\xi\int_0^{\xi} d\eta\int_0^{\eta}\frac{\cos\operatorname{arctg}(b'\xi+\tau_0)}{1-a\xi}\,d\xi;\end{aligned}\right\}\tag{1.40}$$

$$\left.\begin{aligned}&\int_0^{t_f} d\xi\int_0^{\xi} y\,dt=n_0g_0t_f^4I_7,\\&I_7=\int_0^1 d\xi\int_0^{\xi} d\eta\int_0^{\eta} d\mu\int_0^{\mu}\frac{\sin\operatorname{arctg}(b'\xi+\tau_0)}{1-a\xi}\,d\xi;\end{aligned}\right\}\tag{1.41}$$

$$\left.\begin{aligned}&\int_0^{t_f} d\xi\int_0^{\xi} x\,dt=n_0g_0t_f^4I_8,\\&I_8=\int_0^1 d\xi\int_0^{\xi} d\eta\int_0^{\eta} d\mu\int_0^{\mu}\frac{\cos\operatorname{arctg}(b'\xi+\tau_0)}{1-a\xi}\,d\xi.\end{aligned}\right\}\tag{1.42}$$

Here: a is the fuel filling factor ($a = G_{mk}/G_{01}$, where G_{01} and m_1 is the initial weight and the weight of fuel in the second stage), and
$b' = bt_f$ (the stroke is omitted in the following).

Let us rewrite Eqs. (1.133) with reference to the expressions obtained for the integrals $I_1, \ldots, I_8$:

1.
$$\tau_0 = \frac{w_0}{u_0} = \frac{-n_0 g_0 t_f I_1 + g_0 t_f + w_f - \Delta w}{-n_0 g_0 t_f I_2 + u_f - \Delta u};$$
$$v_0^2 = w_0^2 + u_0^2;$$

2.
$$\Delta w = 2\nu^2 \int_0^{t_f} y\,dt = 2\nu^2 \left[y_0 t_f + w_0 \frac{t_f^2}{2} - g_0 \frac{t_f^3}{6} + n_0 g_0 t_f^3 I_5 \right];$$
$$\Delta u = -\nu^2 \int_0^{t_f} x\,dt = -\nu^2 \left[x_0 t_f + u_0 \frac{t_f}{2} + n_0 g_0 t_f^3 I_6 \right];$$

3.
$$\tilde{x}_f = x_0 + u_0 t_f + n_0 g_0 t_f^2 I_4;$$
$$\tilde{y}_f = y_0 + w_0 t_f + n_0 g_0 t_f^2 I_3 - g_0 \frac{t_f^2}{2};$$

4. /71
$$x_f = x_0' + u_0' t_f + n_0 g_0 t_f^2 I_4 + \Delta x;$$
$$y_f = y_0' + w_0' t_f + n_0 g_0 t_f^2 I_3 - g_0 \frac{t_f^2}{2} + \Delta y;$$
$$\Delta x = -\nu^2 \left[x_0 \frac{t_f^2}{2} + u_0 \frac{t_f^3}{6} + n_0 g_0 t_f^4 I_8 \right];$$
$$\Delta y = 2\nu^2 \left[y_0 \frac{t_f^2}{2} + w_0 \frac{t_f^3}{6} - g_0 \frac{t_f^4}{24} + n_0 g_0 t_f^4 I_7 \right];$$

5.
$$w_f = -u_f \frac{x_f}{R + y_f};$$

6.
$$h_f = y_f + \frac{x_f^2}{2R}.$$

The initial point (x_0, y_0) must be selected beyond the limits of the dense atmospheric layers, since as indicated above, the program of the powered section is selected in the dense layers independently of the solution of the variational problem.

This method of calculation can be used not only for two-stage, but even three-stage rockets. By a stage we mean sections of a